

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 2, 2003, 04:55:28 ; Search time 272 Seconds
(without alignments)
822.725 Million cell updates/sec

Title: US-09-977-260-2
Perfect score: 2671
Sequence: 1 MACRGLSVWRAFHGCDSEAE.....PASVSGDADGSTRSPRQEP 507

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 569978 seqs, 220691566 residues
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xip
-O=/cgn2_1/USPTO.spool/US09977260/rumat_29072003_092741_14649/app_query.fasta.1.647
-DB=issued.patents.NA -Qfmt=fastap -SUFFIX=ini -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USPR=US09977260.ecgn_1.1.85.gunat_29072003.092741_14649 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents.NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCRTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2671	100.0	1942	2	US-08-604-989A-11
2	2671	100.0	2000	4	US-08-426-509A-11
3	2671	100.0	2000	4	US-08-232-545-1
4	2671	100.0	2000	5	PCT-US95-05008-1
5	2662	99.7	1521	2	US-08-604-989A-10
6	2591.5	97.0	1987	2	US-08-876-882-1
7	2591.5	97.0	1987	4	US-09-315-928-1
8	2444	91.5	1398	2	US-08-604-989A-9
9	2020.5	75.6	1713	4	US-09-741-154-1
10	1469.5	55.0	16389	4	US-09-741-154-3
11	1269	47.5	738	2	US-08-604-989A-8
12	737.5	27.6	1602	1	US-07-820-011A-1

13	737.5	27.6	1602	5	PCT-US93-00445-1	Sequence 1, Appl1
14	727	27.2	1611	1	US-07-820-011A-3	Sequence 3, Appl1
15	727	27.2	1611	5	PCT-US93-00445-3	Sequence 3, Appl1
16	727	27.2	2129	4	US-09-016-434-1452	Sequence 1452, Ap
17	710	26.6	4517	5	PCT-US93-06251-83	Sequence 83, Appl
18	702.5	26.3	2647	4	US-09-220-132-77	Sequence 77, Appl
19	702.5	26.3	2647	5	PCT-US93-06251-77	Sequence 77, Appl
20	691.5	25.9	1491	3	US-09-006-675-1	Sequence 1, Appl1
21	691.5	25.9	1491	3	US-09-228-603A-11	Sequence 1, Appl1
22	683.5	25.6	3623	1	US-08-306-691B-35	Sequence 35, Appl
23	680	25.5	1467	4	US-09-579-182-2	Sequence 2, Appl1
24	680	25.5	1548	4	US-09-099-053-1	Sequence 1, Appl1
25	656	24.6	2827	4	US-08-492-723-1	Sequence 1, Appl1
26	654.5	24.5	2770	4	US-08-426-509A-5	Sequence 5, Appl1
27	654.5	24.5	2770	5	US-08-232-545-5	Sequence 5, Appl1
28	654.5	24.5	2770	5	PCT-US95-05008-5	Sequence 5, Appl1
29	654.5	24.5	7607	1	US-08-222-616-19	Sequence 19, Appl
30	654.5	24.5	7607	4	US-08-446-648-19	Sequence 19, Appl
31	654.5	24.5	7607	5	PCT-US95-04228-19	Sequence 19, Appl
32	608.5	22.8	2674	4	US-09-817-180-1	Sequence 1, Appl1
33	607	22.7	2505	1	US-08-391-615-1	Sequence 1, Appl1
34	595.5	22.3	3503	1	US-07-631-717A-1	Sequence 1, Appl1
35	595.5	22.3	3503	1	US-08-166-717D-1	Sequence 1, Appl1
36	591	22.1	1875	1	US-08-070-165F-3	Sequence 3, Appl1
37	591	22.1	1875	2	US-08-885-418-3	Sequence 3, Appl1
38	587	22.0	2456	4	US-09-016-434-1476	Sequence 1476, Ap
39	587	22.0	2500	4	US-08-426-509A-3	Sequence 3, Appl1
40	587	22.0	2500	4	US-08-232-545-3	Sequence 3, Appl1
41	587	22.0	2500	4	PCT-US95-05008-3	Sequence 3, Appl1
42	582	21.8	2049	4	US-09-099-749-10	Sequence 10, Appl
43	581.5	21.8	5993	3	US-09-383-630-1	Sequence 1, Appl1
44	581.5	21.8	5993	3	US-09-383-630-2	Sequence 2, Appl1
45	576.5	21.6	2469	1	US-07-997-133-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-604-989A-11
Sequence 11, Application US/08604989A
Patent No. 5834208
GENERAL INFORMATION:
APPLICANT: Sakano, S.
TITLE OF INVENTION: NO. 5834208el Tyrosine Kinase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/604,989A
FILING DATE: February 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Charles F. Miller
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 1920-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1942 base pairs
TYPE: nucleic acid

```

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: human
STRAIN: UT-7
US-08-604-989A-11

Alignment Scores:
Pred. No.: 2,17e-257 Length: 1942
Score: 2671.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-977-260-2 (1-507) x US-08-604-989A-11 (1-1942)
QY 1 MetAlaGlyArgGlySerLeuValSerTTPAlaGAlaPheHisGlyCysAspSerAlaGlu 20
Db ATGGCGGGGAGAGGCTCTGTTCCTGGCGGGCAATTCACGGCTGTGATCTGCTGAG 267
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTPrHisProProProValSerAla 40
Db GAACTTCCCGGGGTGAGCCCCCGCTTCTCCAGCTGGCACCCCTCCGCTCTAGCC 327
QY 41 ArgMetProThrArgArgTPrAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
Db AGGATCCCAACGAGCGCTGGCGCCCGGCGCACCTGATACCAAAATGCCAGCACCC 387
QY 61 ArgProLysProGlyLysLeuAlaPheArgLysGlyAspValAlaThrIleLeuGluAla 80
Db CGGCCCAAGCCAGGGAGCTGGCTTCCGCAAGGGGAGCGTGGTCAACATCTGGAGGCC 447
QY 81 CysGluAsnLysSerTTPArgValLysHisHisThrSerGlyGlnGluGlyLeuLeu 100
Db TGGGACACACAGAGCTGTACCGCGCTCAAGCACACACAGTGAAGAGGAGGCTGCTG 507
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
Db GAGGCTGGGGCGGTGGGAGCGGGAGCGGCTCTCCGACAGCCCAAGCTCAAGCTCATG 567
QY 121 ProTPrPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
Db CCGTGGTTCACGCGGAGATCTCGGGCCAGGAGGCTGTCCAGAGTGGACCTCCCGAG 627
QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuCysVal 160
Db GATGGGCTGTCTCTGGTGGGAGATCTCGGCGCCACCCGGGAGACTACGTCTGTCTG 687
QY 161 SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle 180
Db AGCTTTGGCGCGGACGTATCCACATCCCGCTGTCTCACCGGAGCGCCACCTCACATTC 747
QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLysAspLys 200
Db GATGAGCGCGGTCTTCTGTGCAACCTCAAGATGATGTGAGATTAACAGCAAGCAAG 807
QY 201 GlnAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
Db GGGCGATCTGCACCAAGCTGTGTGACACCAAAAGCGGAAACACGGGACCAAGTGGCGG 867
QY 221 GlnGluLeuAlaArgAlaGlyTPrLeuLeuAsnLeuGlnHisLeuThrLeuGlnAlaGln 240
Db GAGGAGCTGGCCAGGCGGGGTGGTTAAGAACCTGCACATTTGACATTTGGAGACAG 927
QY 241 IleGlyGlnGlyGluPheGlyAlaValIleuGlnGlyLysIleuGlnGlyLysValAla 260
Db ATGGGAGGAGGAGAGTTGTGAGCTGTCTGCAAGGGTGAATACCTGGGCAAAAGGGGCC 987
QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
Db GTGAAGATATATCAAGTGTATGTGACAGCCAGGCTTCTGTGAGCAGACAGCGGCTCATG 1047

```

```

QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
Db ACAGAGATGACACACAGAGAACTGTGTGCTCTGGGGGTGATCTCTGCACACGAGGCGTG 1107
QY 301 TyrTlleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgTPrArgGly 320
Db TACATTGTGATGAGACGATGAGCAAGGGCAACCTGGTGAACCTTGTGGGAGCCGGGGGT 1167
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlnGlyMet 340
Db CGAGCCCTGTGTACACACCGCTCAAGCTCTGTGAGATTCTTCTGTGACGTGGCCAGGCAAG 1227
QY 341 GlnTyrLeuGlnSerLysLysLeuValHisArgAspLeuAlaIleArgAsnIleLeuVal 360
Db GAGTACTGTGAGAGCAAGAACTTGTGACACCGGACCTGGCCCGCCGCAACATCTGTGTC 1287
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
Db TCAGAGAGCTGTGTGGCCAAAGGTCAAGCTGACTTGTGGCTGGCCAAAGCCAGAGGAGG 1347
QY 381 LeuAspSerSerArgLeuProValLysTPrThrAlaProGluAlaLeuLysHisGlyLys 400
Db CTGAGACTCAAGCCGCTGCCCTCAAGTGTGACGCGCCCGAGGCTCTCAACACGAGGAAAG 1407
QY 401 PheThrSerLysSerAspValTyrSerPheGlyValLeuLeuTPrGluValPheSerTyr 420
Db TTCACACAGCAAGCTGATGTCTGAGTTTGGGGTCTCTCTGGGAGGCTCTTCATATAT 1467
QY 421 GlyArgAlaProThrProLysMetSerLysGlnValSerGluAlaValGlnGlyLysGly 440
Db GAGCGGGCTCGTACCTTAATGCTCAAGAGAGTGTGCGAGGCGCGTGGAGAAAGGGG 1527
QY 441 TyrArgMetGlnProProGluGlyCysProGlyProValHisValLeuMetSerSerCys 460
Db TACCGATGGAACCCCGCGAGGCTGTCCAGGCGCGCTGTCTCATATGACACATCTG 1587
QY 461 TPrGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
Db TGGGAGGAGAGAGCCCGCCGCGCCAGCCCTTCCGAAACTGGCCAGAAAGCTGGCGCCGG 1647
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
Db GAGCTAGGAGTGTGAGTGTGCGCCAGCTGTCTGAGGAGAGAGCCGACGAGCTTCACCC 1707
QY 501 SerProArgSerGlnGluPro 507
Db TCGCCCCGAAAGCCAGAGAGCC 1728

RESULT 2
US-08-426-509A-1
Sequence 1, Application US/08426509A
Patent No. 6326469
GENERAL INFORMATION:
APPLICANT: Ulirich, Axel
APPLICANT: Gishitsky, Mikhail
APPLICANT: Sures, Irman G.
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN
TITLE OF INVENTION: TYROSINE KINASES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York,
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/426,509A
 FILING DATE: 21-APR-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/232,545
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: COITZEL, Laura A
 REGISTRATION NUMBER: 30,742
 REFERENCE/DOCKET NUMBER: 7683-0074-999
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-869-9741
 TELETYPE: 66141 PENNTE
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2000 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 US-08-426-509A-1

Alignment Scores:
 Pred. No.: 2,27e-257 Length: 2000
 Score: 2671.00 Matches: 507
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-977-260-2 (1-507) x US-08-426-509A-1 (1-2000)

QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
 Db 258 ATGGGGGGGCGAGGCTCTGCTGCTGCGGGGCGATTTACAGCGCTGATTCGCTGAG 317
 QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
 Db 318 GAACCTCCCGGGGTAGCCCGGCTTCTCCGAGGCTGAGCCCGCTCCGCTCACCC 377
 QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrIleCysGlnHisThr 60
 Db 378 AGGATGCCACAGAGCGCTGGGCGCCCGGCAACCGATTCACCAATGGAGACACC 437
 QY 61 ArgProLysProGlyGluLeuAlaPheArgGlyGlyAspValValThrIleLeuGluAla 80
 Db 438 CGCCCAAGCCAGGGAGGAGCTGGCTCCGCAAGGCGAGCTGATCACCATCCTGGAGGCC 497
 QY 81 CysGluAsnLysSerTrpTrpArgValIleHisThrHisThrSerGlyGlnGluGlyLeu 100
 Db 498 TGGCAGACAGAGAGTGTACCGGCTCAACACACACAGAGTGGACAGAGGGGCTGTG 557
 QY 101 AlaAlaGlyAlaLeuAlaGlyGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 Db 558 GCACCTGGGGCGCGGCGGAGCGGAGGCGCTTCGCGAGACCCCAAGCTCAGCCTCATG 617
 QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
 Db 618 CCGGTGTTCCACGGGAGAGATCTGGGAGGAGGCTGCGCAGACGTCGACCTCCGAG 677
 QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysVal 160
 Db 678 GATGGGCTGTTCGTGGTGGGAGATCCGCGCGGCAACCGCGAGTACTGCTGTGCTG 737
 QY 161 SerPheGlyArgAspValIleHisTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
 Db 738 AGCTTTGGCCGCGAGCATCATCCATCCGCTGCTGACCGCGAGCGCACCTCATCATC 797
 QY 181 AspGluAlaValIlePhePheCysAsnLeuMetAspMetValGlnHisTrpSerTrpLysAspLys 200
 Db 798 GATAGGCGCGTGTCTTCTGCAACCTATGACATGTGGTGGACATTTACAGCAAGACAG 857
 QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220

Db 858 GCGGCTATGTGCACCAAGCTGTGAGACCAAAAGCGGAAACACAGGACCAAGTCGGCGAG 917
 QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
 Db 918 GAGAGAGCTGCCAGGGCGGGTGTACTGAACTGACACATTTGATTTGGAGACAG 977
 QY 241 IleGlyGlnGlyLysPheGlyAlaValLeuGlnGlnGlyGlyTrpLeuGlyGlnLysValAla 260
 Db 978 ATCGGAGAGGAGAGATTTGGAGCTGCTGCGAGGGTGAATACCTGGGCGAAAAGGTGGCC 1037
 QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGlyThrAlaValMet 280
 Db 1038 GTGAAGAATATCAATGTGATGTGACAGCCAGGCTTCTGAGAGAGAGCCCTCATG 1097
 QY 281 ThrLysMetGlnHisGlnAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
 Db 1098 ACGAAGATGCACACAGAGAACCTGGTGGCTCTGCGGCGTATCTGCACACAGGGGCTG 1157
 QY 301 TyrIleValIleMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
 Db 1158 TACATTTCATGAGACAGCTGAGCAAGGCAACCTGTGAACCTTCTCGGACCGGGGT 1217
 QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMet 340
 Db 1218 CGAGCCCTCGTGAACACCGCTGAGCTCTGAGTTTCTGACAGTGGCCGAGGGCATG 1277
 QY 341 GluTrpLeuGluSerLysLysLeuValHisArgAspLeuAlaIleArgAsnIleLeuVal 360
 Db 1278 GAGTACCTGGAGAGCAAAACCTTGTGACCGCGACCTGCGGCGCCCGCAACATCTCGTGC 1337
 QY 361 SerGluAspLeuValAlaIleValSerAspPheGlyLeuAlaLysAlaGluArgGly 380
 Db 1338 TCAGAGACACTGTGTGCCAAGTCAAGCACTTTGGCTTGGCCAAAGCCGAGCGAGGGG 1397
 QY 381 LeuAspSerSerArgLeuProValIleTrpThrAlaProGluAlaLeuLysHisGlyLys 400
 Db 1398 CTAGACTCAACCGGCTCGCTCGTCAAGTGGAGCGGCGCGAGGCTCTCAACACGGGAG 1457
 QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrp 420
 Db 1458 TTCACACAGAGTGGAGTGTGAGAGTTTGGGTGCTGCTGTGGAGCTTCTCATAT 1517
 QY 421 GlyArgAlaProTrpProLysMetSerLeuLysGluValSerGluAlaValGluGly 440
 Db 1518 GGAGGGGCTCGTACCTCAAAATGTCACTGAAAGAGTGTGGAGCGGTGGAGAAAGGGG 1577
 QY 441 TyrArgMetGluProProGlnGlyCysProGlyProValHisValLeuMetSerCys 460
 Db 1578 TACCGCATGAAACCCCGGAGGCTGTCCAGGCGCCGTCACGTCTCATAGACAGCTGC 1637
 QY 461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
 Db 1638 TGGAGAGCAGAGCCCGCCCGCCGCAACCTTCCGCAAACTGGCGGAGAAAGTGGCCGG 1697
 QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
 Db 1698 GAGCTACGAGTGTGAGGTGCCACAGCTCTCGTTCAGGGGAGAGCGGAGGCTCCAC 1757
 QY 501 SerProArgSerGlnGluPro 507
 Db 1758 TCGCCCGGAAGCCAGAGGCC 1778

RESULT 3
 US-08-232-545-1
 : Sequence 1, Application US/08232545
 : Patent No. 6506578
 : GENERAL INFORMATION:
 : APPLICANT: Ullrich, Axel
 : APPLICANT: Gishizsky, Mikhail
 : APPLICANT: Sures, Irmann G.
 : TITLE OF INVENTION: No. 6506578el Megakaryocytic Protein Tyrosine
 : TITLE OF INVENTION: Kinases

NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,545
FILING DATE: 22-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2000 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-232-545-1

Alignment Scores:
Pred. No.: 2,27e-257 Length: 2000
Score: 2671.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-977-260-2 (1-507) x US-08-232-545-1 (1-2000)

QY 1 MetalGlyArgGlySerLeuValSerTPArgAlaPheHisGlyCysAspSeraIaGlu 20
DB 258 ATGGCGGGGCGAGGCTCTCTGTTCTCTGCGGGCATTTCCAGCGTGATTCGTGAG 317
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProProValSeraIa 40
DB 318 GAACCTTCCCGGGGTAGCCCCCGCTTCCTCGAGCTGGCACCCCTCCCGCTCAGCC 377
QY 41 ArgMetProThrArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
DB 378 AGGATCCCAAGGCGGTGGCGCCCGGCGCACCATGTATCCAAATGGCGACACC 437
QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaThrIleLeuGluAla 80
DB 438 CGCCCCAAGCCAGGGGAGCTGGCTCCGCAAGGGCGAGTGTGCACCATCTGGAGGCC 497
QY 81 CysGluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGluLeuLeu 100
DB 498 TGGGAGACAAAGAGCTGTACCGCGTCAAGCAACACACAGTGGACAGGAGGGGTG 557
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
DB 558 GCAAGTGGGGCGGTGGGAGCGGAGGCCCTCTCCGCGAGACCCCAAGCTCAGCTCATG 617
QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
DB 618 CAGTGTGTCACGAGGAAATCTCGGGCCAGAGGCTGTCCAGAGCTGCAGGCTCCCGAG 677
QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysVal 160

DB 678 GATGGCTGTTCTGGCGGAGTCGCGGCCACCCGCGACATCAGCTGTGCTG 737
QY 161 SerPheGlyArgAspValIleHisTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
DB 738 AGCTTTGGCGCGCACATCATCCATCCGCGGTCTGACCGGACCGCCACCTCATATC 797
QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTrpSerLysAspLys 200
DB 798 GATGAGCGCGTGTCTTCTTCACATCATGATGATGGATGAGCATTAACAGCAAGCAAG 857
QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSeraIaGlu 220
DB 858 GCGCGTATCTGCACCAAGCTGTGAGAACCAACCGGAAACCGGACCAAGTCGGCCGAG 917
QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
DB 918 GAGGAGCTGCCCGGCGGGCGGTGTACTGTAACCTGACGATTTGACATTTGGAGCAG 977
QY 241 IleGlyGluGlyLupPheGlyAlaValLeuGlnGlyLupTrpLeuGlyGlnLysValAla 260
DB 978 ATCGAGAGAGGAGATTGGAGCTGTCTGACAGGTGAGTACCTGGGCAAAAGTGGGCC 1037
QY 261 ValLysAsnLleLysCysAspValIThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
DB 1038 GTGAAGAATATCAAGTGTGTGACAGCCGACCTTCTCTGACAGAGAGCGCCGTCATG 1097
QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
DB 1098 ACGAAGATGCACACAGAACCTGCTGCTCTCTGGCGTGTATCTTGACACAGGGGCTG 1157
QY 301 TyrIleValMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgTrpArgGly 320
DB 1158 TACATGTGATGAGACAGTGAAGCAAGGCAACCTGTGAACCTTCTGGGAGACCGGGGT 1217
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMet 340
DB 1218 CGAGCCCTGTGAAACACGCTGCTGCTGCAATTTCTCTGACCTGGCCGCGGCGCATG 1277
QY 341 GluTrpLeuGluSerLysLysLeuValHisArgAspLeuAlaIlaArgAsnIleLeuVal 360
DB 1278 GAGTACTGAGAGCAAGAAAGCTGTGCAACCGGAGCTGCGCCCGCCGCAACATCTGCTG 1337
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
DB 1338 TCAGAGAGACTGTGTGGCCAAAGTCAAGCATTTGGCTGGCCAAACCGGAGGAGGG 1397
QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
DB 1398 CTAGACTCAAGCGGCTGCCGTCAGTGAAGTGAAGCGCCCGGAGGCTCTCAAAACCGGGAAG 1457
QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrp 420
DB 1458 TTCACAGCAAGTGGATGTGTGAGTTTGGGGTCTCTCTGAGAGGTCTTCTCATAT 1517
QY 421 GlyArgAlaProTrpProLysMetSerLeuLysGlyValSerGluAlaValGluLysGly 440
DB 1518 GGAGCGGCTCCGACCTTAAATGTCACTGAAGAGAGTCTCGAGGCGCTGAGAAAGGGG 1577
QY 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerCys 460
DB 1578 TACCGCATGAAACCCCGGAGGGCTGTCCAGGCGCCGTCACTCTCTTGAAGCAAGTGC 1637
QY 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
DB 1638 TGGGAGGCGAGACCCCGCGCGGCGCACCTTCCGAAACTGGCGAAGAGCTGGCGCCG 1697
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspLysSerThr 500
DB 1698 GAGCTACGAGTGCAGAGTCCCGACGCTCGTCTGAGGACAGAGCGCCGACGGCTCACCC 1757
QY 501 SerProArgSerGlnLupPro 507
|||||

Db 1758 TCGCCCGAAGCAGGAGCC 1778

RESULT 4
PCT-US95-05008-1
Sequence 1, Application PC/TUS9505008
GENERAL INFORMATION:
APPLICANT: Sugen, Inc.
APPLICANT: 515 Galveston Drive
APPLICANT: Redwood City, California 94063-4720
APPLICANT: United States of America
APPLICANT: Missionschaften E.V.
APPLICANT: Hofgarten Str. 2
APPLICANT: Munchen 80539
APPLICANT: Germany
TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
NUMBER OF INVENTION: Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penile & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05008
FILING DATE: 24-APR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,545
FILING DATE: 22-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-074
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNITE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2000 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
PCT-US95-05008-1

Alignment Scores:
Pred. No.: 2 276-257 Length: 2000
Score: 2671.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: Indels: 0
DB: Gaps: 0

US-09-977-260-2 (1-507) x PCT-US95-05008-1 (1-2000)

QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
DB 258 ATGGGGGGGCGAGGCTCTCTGTTCTCGCGGCGCATTTCAAGGCTGTGATTTCTGCTAG 317
QY 21 GluLeuProAlaValSerProAlaGlyPheLeuAlaTrpHisProProValSerAla 40
DB 318 GAACCTCCCGGGTACGCCCCGCTTCTCCGAGCCTGCACCCCGCTCCGCTCACACC 377
QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGlnHisThr 60

Db 378 AGGATGCCAAGCAGGGGCTGGGCCCCGGGCGACCCAGTGTATCCAAATGCCAGCACACC 437
QY 61 ArgProLysProGlyGluLeuAlaPheArgGlySGlyAspValAlaThrIleLeuGluAla 80
DB 438 CGCCCCAAGCCAGGGAGGCTGGCTTCCCAAGGGCGAGCTGTCACATCTGGAGGCC 497
QY 81 CysGluAsnLysSerTrpTrpArgValLysHisIleThrSerGlyGlnGluLeu 100
DB 498 TCGCAACACAGAGCTGGTACCGCGTCAACACACACACAGGAGAGGGGCTGTG 557
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
DB 558 GCAGCTGGGGGCGCTGGGGAGGGAGGGCGCCCTCCGCGACCCCAAGCTCAGCTCAG 617
QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaGlnGluLeuInProGlu 140
DB 618 CCGTGGTTCACAGGAGATCTCGGGCCAGGAGGCTGTCCACAGCTGCAGCTCCGAG 677
QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuGlyVal 160
DB 678 GATGGGCTTCTCTGGTGGGAGTCCCGCGCCACCCCGGACTACGTCCTGTGCGTG 737
QY 161 SerPheGlyArgAspValIleHisTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
DB 738 ACCTTGGCGCGAGCGTATCCACTACCGCGTGTGACACCGGAGGGCCACTCACAATC 797
QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTrpSerLysAspLys 200
DB 798 GATGAGGCGCTGTCTTCTGCAACCTCATGACATCGTGACATTCAGCAGGAGACAAG 857
QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
DB 858 GCGGCTATCTGCACCAAGCTGTGAGACCAAGCGGAAACAGGAGCAAGTCCGCGAG 917
QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGluAlaGln 240
DB 918 GAGGAGCTGGCGAGGGGCGGCTGTACTGAGACCTGCAGCATTTGAGATTGGAGCACAG 977
QY 241 IleGlyGluGlyGluPheGlyValAlaLeuGlnGlnGlyGlyTrpLeuGlyGlnLysValAla 260
DB 978 ATCGGAGAGGAGAGATTTGGAGCTGTCTGCGAGGCTGTGATACCTTGGGCAAAAGTGGCC 1037
QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGlyThrAlaValMet 280
DB 1038 GTGAGGATATTAATGTGATGTGACACCCAGGCTTCTGTGACGAGAGGCGCTCATG 1097
QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
DB 1098 ACGAAGATGCACACAGACAGACCTGTGCGTCTCCGTGGCGGTGATCTGCACAGGGGCTG 1157
QY 301 TyrIleValMetGlnHisValSerLysGluAsnLeuValAsnPheLeuAlaArgTrpGly 320
DB 1158 TACATTGTCAATGAGACAGTGTGACAGGCAACCTGTGTAACTTTCTCGGAGCGGGGT 1217
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMet 340
DB 1218 CGAGCCCTCGTGAACACCGCTCAGCTGTGACAGTTTCTGTGACAGTGGCCGAGGGCATG 1277
QY 341 GluTrpLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
DB 1278 GAGTACCTGGAGAGACAGAACCTTGTGCACGCGCACTGGCGCGCCGCAACATCTGTGTC 1337
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
DB 1338 TCAAGAGACTGTGTGCGCAAGAGTCAAGCACTTTGTGGCTGGCGCAAGCGGAGGAGGG 1397
QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
DB 1398 CTAGACTAACCGGCTCCGCTCAAGTGAAGAGCGCGCCGAGGCTCTAAACACAGGGAG 1457
QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrp 420

Db 1458 TTCACGACGATGCGATGCTGGAGTTTGGGCTGCTCTCTGGAGGCTTCTCATAT 1517
 QY 421 GYARGLAARProTyProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
 Db 1518 GGAGGGGCTCCGCTACCTAAATGTCACTGAAGAGGTGTCGAGGGCGTGGAGAGGGG 1577
 QY 441 TYTRGmetGluProProGluGlyCysProGlyProValHisValLeuMetSerCys 460
 Db 1578 TACCGCATGAAACCCCGAGGGGCTGTCAGGCCCGCTGCACGCTCATGACGACGCTGC 1637
 QY 461 TTPGLuAgluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
 Db 1638 TGGGAGGACAGAGCCCGCCCGCCGACCTTCGCAAACTGGCCGAGAGAGTGGCCGG 1697
 QY 481 GluLeuArgSerAlaGluAlaProAlaSerValSerLysGluAlaAspLysThr 500
 Db 1698 GAGCTACGACAGTGCAGGTGCCCCAGCTCGTTCAGGGCAGAGCCGAGGCTCCACC 1757
 QY 501 SerProArgSerGlnGluPro 507
 Db 1758 TCGCCCGAAGCCGAGAGCCC 1778

RESULT 5

US-08-604-989A-10
 : Sequence 10, Application US/08604989A
 : Patent No. 5834208

GENERAL INFORMATION:

APPLICANT: Sakano, S.
 TITLE OF INVENTION: No. 5834208a1 Tyrosine Kinase
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds LLP
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/604,989A
 FILING DATE: February 23, 1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Charles E. Miller
 REGISTRATION NUMBER: 24,576
 REFERENCE/DOCKET NUMBER: 1920-026
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1521 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 ORIGINAL SOURCE:
 ORGANISM: human
 STRAIN: UF-7
 US-08-604-989A-10

Alignment Scores:

Pred. No.: 1.17e-256
 Score: 2662.00
 Percent Similarity: 99.808
 Best Local Similarity: 99.808
 Query Match: 99.668
 DB: 2
 Gaps: 0

US-09-977-260-2 (1-507) x US-08-604-989A-10 (1-1521)
 QY 1 MetLlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
 Db 1 ATGGGGGGGAGGCTCTGTGGTTCTCTGGGGGACATTCACGGCTGTGATTCGCTGAG 60
 QY 21 GluLeuProArgValSerProArgPheLeuArgLalaTrpHisPropProValSerAla 40
 Db 61 GAATTCCTCCGGGTGAGCCCTCCGCTTCGAGCCCTGAGCCCTCCGCTCCGCTCCGCC 120
 QY 41 ArgMetProThrArgArgTrpAlaProGlyTrpGlnCysIleThrLysCysGlnHisThr 60
 Db 121 AGGATGCCAACGAGGCGCTGGCCCGCCGAGCCACCGAGTATACCAAAATCGAGACACC 180
 QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaThrIleLeuGluAla 80
 Db 181 CGCCCCAGCCAGGGGAGCTGCTCCGCAAGGGGCGAGCTGTACCATCTGGAGGCC 240
 QY 81 CysGluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGluLysLeu 100
 Db 241 TCGAGAAACAAGAGCTGTACCGGTCACACACACACAGTGAGAGAGGGGCTGCTG 300
 QY 101 AlAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 Db 301 GCACTGGGGCGCTGGCGGAGCGGAGGCCCTCTCCGACAGCCCAAGCTCACCTCATG 360
 QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
 Db 361 CCGTGTGTCACGAGAAATCTGGGCGAGGAGGCTGTCCAGACCTGACGCTCCCGAG 420
 QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysVal 160
 Db 421 GATGGCTGTCTCTGGTGGGAGGAGTCCGCGCCACCCCGGACATACGCTGTGCGTG 480
 QY 161 SerPheGlyArgAspValIleHisTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
 Db 481 AGCTTTGGCCGCGACGATCATCACTACCGGTCTGTGAGCCGACGCGCACCTCACAAATC 540
 QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTrpSerLysAspLys 200
 Db 541 GATGAGGCGCGTGTCTTCTTCACACCTCATGAGCATGTGAGCATTCACAGAGCAAG 600
 QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
 Db 601 GCGCTATCTGCACCAAGCTGTGAGACCAACGAGAAACGAGGACCAAGTGGCGGAG 660
 QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
 Db 661 GAGGAGCTGGCCAGGGCGGGCTGTACTGAACTGCACATTGTGACATTGGGAGACAG 720
 QY 241 IleGlyGluGlyGluPheGlyAlaValLeuGlnGlyGlyTrpLeuGlyGlnLysValAla 260
 Db 721 ATCGAGAGGAGGAGATTGGAGCTGTCTGACAGGAGTGTGAGTGGGCAAAAGTGGCC 780
 QY 261 ValLysAsnLysLysCysAspValIleAlaGlnAlaPheLeuAspLysThrAlaValMet 280
 Db 781 GTGAAGAATAATCAAGTGTATGTGACAGCCAGGCTTCTGTGACAGAGAGGCGCTCATG 840
 QY 281 ThrLysMetGlnHisGlnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
 Db 841 AGCAAAATGCAACACAGAACTGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 QY 301 TyrIleValMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgTrpArgGly 320
 Db 901 TACATGTGATGAGACAGCAGAGGAGGAGCACTGTGACATTCTGGGAGCCCGGGGT 960
 QY 321 ArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGluGlyMet 340
 Db 961 CGAGCCCTGTGTAAACCCCTCAGCTCCGAGTTTCTCTGACGCTGCGCAGGCGCATG 1020
 QY 341 GluTrpLeuGlnSerLysLysLeuValHisArgAspLeuAlaArgAsnIleLeuVal 360

Db 1021 GAGTACTGGAGACCAAGACCTTGTCACCGCAGCCTGGCCGCCGCCACATCTCGTGC 1080
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGlyArgLysGly 380
Db 1081 TCAGAGACCTGTGGCCAGAGCTGACACTTTGGCCCTGGCCAAAGCCGAGCGAGGGG 1140
QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
Db 1141 CTAGACTCAAGCCGGCTGCCCCGTCAAGTGAAGGCGCCGAGGCTCTCAACACGCGAG 1200
QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyr 420
Db 1201 TTCACGACGACAGTGGATGCTGTGAGTTTGGGGTGGCTGCTGGAGGCTTCTCATAT 1260
QY 421 GLYArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
Db 1261 GGAGCGGCTCCGTACCTTAATAATGTCACTGAAGAGGTGTGGAGGCCGTGGAGAGGG 1320
QY 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerCys 460
Db 1321 TACCGCATGAACCCCGCAGAGGCTGTCCAGGCCGCCGTCATAGAGCAGCTGC 1380
QY 461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
Db 1381 TGGGAGGACAGCCCGCCGCGCCGACCTTCCGCAAACTGGCCGAGAACCTGGCCGG 1440
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyHisAspAlaAspGlySerThr 500
Db 1441 GAGCTACGCACTGCAAGGTGCCCCAGCTCTCGTCTGAGGCGAGAGGCCGAGGCTCCAC 1500
QY 501 SerProArgSerGlnGluPro 507
Db 1501 TCGCCCGCAAGCCAGAGGCC 1521

RESULT 6
US-08-876-882-1
: Sequence 1, Application US/08876882
: Patent No. 5981201
: GENERAL INFORMATION:
: APPLICANT: Avraham, Hava
: TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hamilton, Brook, Smith & Reynolds P.C.
: STREET: Two Militia Drive
: CITY: Lexington
: STATE: MA
: COUNTRY: USA
: ZIP: 02173-4799
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: Windows
: SOFTWARE: FastSeq for Windows Version 2.0b
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/876,882
: FILING DATE: 16-JUN-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/035,228
: FILING DATE: 08-JAN-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Doreen, Hogle M
: REGISTRATION NUMBER: 36,361
: REFERENCE/DOCKET NUMBER: NEDH97-01PA
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 781-861-6240
: TELEFAX: 781-861-9540
: INFORMATION FOR SEQ ID NO. 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1987 base pairs

: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
US-08-876-882-1

Alignment Scores:
Pred. No.: 2,02e-249 Length: 1987
Score: 2591.50 Matches: 502
Percent Similarity: 99.21% Conservative: 1
Best Local Similarity: 99.01% Mismatches: 3
Query Match: 97.02% Indels: 1
Gaps: 3

US-09-977-260-2 (1-507) x US-08-876-882-1 (1-1987)

QY 1 MetaLagIyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
Db 263 ATGGCGGGGCGAGGCTCTGTGTTCTTGGCGGCGATTTACAGGCTGTGATTCGTGAG 322
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProProValSerAla 40
Db 323 GAACCTCCCGGGGTAGCCCGCCCTTCCCGAGCCTTGACACCCCTCCGCTCAGCC 382
QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
Db 383 AGGATGCCAAGAGGCGCTGGGCCCGGCGACCCAGTGTATCCAAATTCGAGACAC 442
QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaThrTleuGluAla 80
Db 443 CGCCCAAGCCAGGGAGCTGGCTCCGCAAGGCGAGCTGTACCATCTCGAGAGGCC 502
QY 81 CysGluAsnLysSerTrpTyrArgValLysHisIsthrSerGlyGlnGlyLeu 100
Db 503 TCGGAGAACAGAGCTGTACCGGTCAAGACACACAGAGTGAGAGAGGCTGTG 562
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
Db 563 GCAGCTGGGGGCTGGGAGGAGGAGGAGGCTCTCCGAGAGCCCAACCTCAGCTCAG 622
QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
Db 623 CCGTGTCCACAGGAGATCTCGGCGCAGAGGCTGTCCAGACCTCAGACCTCCGAG 682
QY 141 AspGlyLeuPheLeuValArgGluSerLarArgHisProGlyAspTyrValLeuGlyVal 160
Db 683 GATGGCTGTCTGTGTGGGAGTCCGCGGCCACCCCGGAGTACGTCTGTGCTG 742
QY 161 SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle 180
Db 743 AGCTTTGGCCCGAGAGTATCCACTACCGGCTGTGACCCGAGCGGCACCTCACAATC 802
QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTyrSerLysAspLys 200
Db 803 GATGAGGCCGTTGTCTTCCCAACCTCATGAGCAGTGTGAGCATTAACAGAGACAA 862
QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAla 220
Db 863 GCGCTATCTCACCAACCTGTGTAGACCAAGCGGAAACACGGGACCAAGTGGCCAG 922
QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
Db 923 GAGAGAGTGGCCAGGGCGGCTGTACTGAACCTGCGACATTTGACATTTGGAGACAG 982
QY 241 IleGlyGluGlyLysPheGlyAlaValAlaLeuGlnGlyLysGlyLysValAla 260
Db 983 ATCGGAGAGGAGAGATTGTGAGCTGTCTCGACAGGTGAGTACCTGGGCGCAAAAGTGGCC 1042
QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLysAspGlyThrAlaValMet 280
Db 1043 GTGAAGATATCAAGTGTGATGTGACACCCAGGCTTCTTGGAGAGACGAGCGCTCATG 1102
QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300

Db 1103 ACGAAGATGCACACGAGAACCTGTGCTCTCTGGGCGGTGATCTGACACAGAGGGCTG 1162
QY 301 TTTTLevalMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
Db 1163 TACATGTGATGAGCAGCAGTGCAGGAGGCAACCTGGTGAACCTTTCGGCGAGCCGGGGT 1222
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlyMet 340
Db 1223 CGAGCCCTCTGGAACACCGCTCACCTCTCTGACAGTTTCTCTGACGCTGGCGAGGGCATG 1282
QY 341 GluTyrLeuGlnSerLysLysValHisArgAspLeuAlaAlaAsnIleLeuVal 360
Db 1283 GAGTACTGAGAGACAAAGACTTGTCTACCGGACCTGGCGCCGCCAATCTTGTGTC 1342
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGlyArgGly 380
Db 1343 TCAGAGAGACCTGTGGCCAGAGTCAGCAGCTTGTGGCTGGCCAAAGCGAGCGAAGGGG 1402
QY 381 LeuAspSerSerArgLeuProValLysTyrThrAlaProGluAlaLeuLysHisGlyLys 400
Db 1403 CTGAGCTCAAGCGCGGTGCTCCGTCAAGTGAAGCGCGCCGAGGCTTCAACACAGGG-- 1459
QY 401 PheThrSerLysSerAspValTyrSerPheGlyValLeuLeuTyrGluValAlaPheSerTyr 420
Db 1460 TTTCACGACGAAAGTCGATGTCTGAGTTTGGGGTGTCTGCTGGGAGGCTTCTCATAT 1519
QY 421 GlyArgAlaProTyrProLysMetSerLeuLysGlyValSerGluAlaValGlyLysGly 440
Db 1520 GGAAGGGCTCCGTACCTAAATGTCACTGAAGAGGTGTGAGGCGCTGGAGAAAGGGG 1579
QY 441 TyrArgMetGluProProGluGlyCysProGlyProAlaHisValLeuMetSerSerCys 460
Db 1580 TACCGCTTGAACCCCGGAGGGGTGTCCAGGCGCCCTGCGACGCTTCAAGAGAGTGTG 1639
QY 461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
Db 1640 TGGGAGGACAGAGCGGCC -CGCGGCGCACCTTCGCGAAACTGGCGAGAAAGCTGGCCGG 1698
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyLysAlaAspGlySerThr 500
Db 1699 GACTACGACAGTGCAGGTGCCCCAGCTCCGTCTCAGGAGCAGAGCGCAGAGTTC -ACC 1757
QY 501 SerProArgSerGlnGluPro 507
Db 1758 TCGCCCCGAGAGCCAGAGGCC 1778

RESULT 7
US-09-315-928-1
; Sequence 1, Application US/09315928
; Patent No. 6368796
; GENERAL INFORMATION:
; APPLICANT: Avraham, Hava
; APPLICANT: Groopman, Jerome E.
; TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT OF
; FILE REFERENCE: NED97-01PAZ
; CURRENT APPLICATION NUMBER: US/09/315,928
; CURRENT FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 08/876,882
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: US 60/035,228
; PRIOR FILING DATE: 1997-01-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1987
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (263)...(1846)
US-09-315-928-1

Alignment Scores:
Pred. No.: 2,02e-249 Length: 1987
Score: 2591.50 Matches: 502
Percent Similarity: 99.21% Conservative: 1
Best Local Similarity: 99.01% Mismatches: 3
Query Match: 97.02% Indels: 3
DB: 4 Gaps: 1

US-09-977-260-2 (1-507) x US-09-315-928-1 (1-1987)

QY 1 MetAlaLysArgGlySerLeuValSerTyrArgAlaPheHisGlyCysAspSerAlaGlu 20
Db 263 ATGGCGGGGAGGCTCTGTGTTTCTGGCGGGCAATTCACGGCTGTGATCTGTGAG 322
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
Db 323 GAACCTCCCGGGGAGACCCCGCTTCTCGAGCTGGACCCCTCCCGTCTGACC 382
QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
Db 383 AGCATGCCAAGAGAGCGCTGGGCCCGCGGACACCAAGTGTATCAACAAATGCGAGACACC 442
QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaTrpIleLeuGluAla 80
Db 443 CGCCCAAGCGAGGGAGCTGGCTTCGCAAGGGCCACGTGTGCACATCTGAGAGCC 502
QY 81 CysGluAsnLysSerTyrTyrArgValLysHisIleThrSerGlyGlnGlyLeu 100
Db 503 TGGGAGAACAAAGCTGTGTACCGCGTCAACACACACACACGATGGACAGAGGGGCTGCTG 562
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
Db 563 GCAGCTGGGGCGCTGGGGGAGCGGGGCCCTCTCCGCAAGCCCAAGCTCACGCTCATG 622
QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
Db 623 CCGTGTTCACGAGGAGATCTGGGCCAGAGAGCTGTCCAGAGCTGACCTCCCGAG 662
QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisIleProGlyAspTyrValLeuCysVal 160
Db 683 GATGGGCTGTCTCGGGCGGAGTCCGGCGGACCCCGGCGACTACGTCTGTGCTG 742
QY 161 SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle 180
Db 743 AGCTTTGGCCGACGATCATCACTACCGGTCTGTGACCGCCACCGCCCTCACATC 802
QY 181 AspGluAlaValAlaPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLysAspLys 200
Db 803 GATGAGCGCTGTCTTCTTCTCAACCTCATGAGATGAGTGAATTCACAGACAGACAG 862
QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
Db 863 GCGGCTATTCGACCAAGCTGTGTGAGACCAAGCGGAAACAGCGGACCAAGTGGCGAG 922
QY 221 GluGluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
Db 923 GAGGAGCTGGCCAGGGCGGGCTGTGTACTGAACCTCAGACATTTGACATTTGGAGACAG 982
QY 241 IleGlyGlnGlyLysPheGlyAlaValIleGlnGlyGluTyrLeuGlyGlnLysValAla 260
Db 983 ATCGGAGAGGAGAGTGTGAGCTGTCTGCAAGGTGAGTGAATCTGGGGCAAAAGGTGGCC 1042
QY 261 ValLysAsnIleLysCysAspValTrpAlaGlnAlaPheLeuAspGluTrpAlaValMet 280
Db 1043 GTGAAGAATATCAAGTGTATGTGACAGCCAGGCTTCTGTGACAGAGAGCGGCTCATG 1102
QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuGlyValIleLeuHisGlnGlyLeu 300
Db 1103 ACGAAGATGCACACAGAAACCTGTGCTCTCTGCGGTGATCTGCAACACAGGGGGCTG 1162
QY 301 TTTTLevalMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320

Db 1163 TACATTGATGAGACAGTGTGCAAGGCAACCTGTGTAAGTCTTCTGGGAGCCCGGGGT 1222
QY 321 ATGAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGluMet 340
Db 1223 CGAGCCCTGCTGAACACCGCTCAGCTCCTGCGAGTTTCTCTCAGGTGGCCGAGGCGATG 1282
QY 341 GUTYrLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
Db 1283 GAGTACCTGTGAGAGCAAGACTTGTGCACCCGACCTGGCCGCCCAACATCTGTGTC 1342
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaAlaGluArgLysGly 380
Db 1343 TCAGGAGACTGTGTGGCCAGAGTGTGCACTTTGGCTGGCCAAAGCCGAGCGAGGAGG 1402
QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuHisGlyLys 400
Db 1403 CTAGACTACAGCCGGCTGCTCCCTCAAGTGAAGGCGCCGAGGCTCTCAACACGG--- 1459
QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyr 420
Db 1460 TTCACACCAAGTGTGATGTCTGAGATTGAGGTGCTGTGAGAGGTCTTCTCATAT 1519
QY 421 GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
Db 1520 GGAGGGCTCGTACCTTAATGTCTACTGAAGAAGGTGTGAGAGCCGTGAGAGAGG 1579
QY 441 TyrArgMetGluProGluGluLysProGluProValHisValLeuMetSerSerCys 460
Db 1580 TACCGCATGGAACCCCGAGAGGCTGTCCAGGCCCGCTGTACGTCTATGACAGCTGC 1639
QY 461 TrpGluAlaGluProAlaArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
Db 1640 TGGGAGGCAAGCCGCC--CGCCGGCCACCTTCCCAAACTGGCGAAGGTGCGCCGG 1698
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
Db 1699 GAGTACCGAGTGTGAGGTGCCACGCTCCGTCTAGGAGGAGCGGACGCTCC-ACC 1757
QY 501 SerProArgSerGlnGluPro 507
Db 1758 TCGCCCGCAGACGACGAGCC 1778

RESULT 8
US-08-604-989A-9
; Sequence 9, Application US/08604989A
; Patent No. 5834208
; GENERAL INFORMATION:
; APPLICANT: Sakano, S.
; TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/604,989A
; FILING DATE: February 23, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles E. Miller
; REGISTRATION NUMBER: 24,576
; REFERENCE/DOCKET NUMBER: 1920-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1398 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: human
; STRAIN: UT-7
; US-08-604-989A-9

Alignment Scores:
Pred. No.: 6,656-235 Length: 1398
Score: 2444.00 Matches: 466
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 91.50% Indels: 0
DB: 2 Gaps: 0

US-09-977-260-2 (1-507) x US-08-604-989A-9 (1-1398)

QY 42 MetProThrArgArgTrpAlaProGluYrGlnCysIleThrLysCysGluHisThrArg 61
Db 1 ATGCCACAGAGCGCTGGGCCCCCGGCAACCGTATCACCAATGGGAGCACACCCGC 60
QY 62 ProLysProGluGluLeuAlaPheArgLysGlyAspValAlaThrIleLeuGluAlaCys 81
Db 61 CCCAAGCCAGAGGAGCTGGCTTCCGCAAGGCGCAGCTGTCCATCCTGTGAGGCTGC 120
QY 82 GluAsnLysSerTrpTyrArgValLysHisThrSerGlyGlnGluLeuAla 101
Db 121 GAGAACAGAGCTGTATCCGCTCAAGCACACACAGTGGACAGAGGGCTGCTGCA 180
QY 102 AlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMetPro 121
Db 181 GCTGGGGCGCTGGGAGCGGAGGAGCGCTTCCGACACCCCAAGCTCAGCTCATGCGG 240
QY 122 TrpPheHisGlyLysIleSerClyGlnGluAlaValGlnGlnLeuGlnProProGluAsp 141
Db 241 TGTTCACAGAGGAAATCTCGGCGCAGAGGCTGTCCAGACCTCAGCTCCAGAGAT 300
QY 142 GlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuGluValSer 161
Db 301 GGGCTGTCTCTGCTGGGAGAGTCCCGGCGCACCCCGGAGCTACTGCTGTGCTGAC 360
QY 162 PheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIleAsp 181
Db 361 TTTGGCCGCGAGCTATCCACTACCGCGTGTGACCGCGAGCGGCGCACCTCAGATCGAT 420
QY 182 GluAlaValPhePheCysAsnLeuMetAspMetValGluHisTyrSerLysAspLysGly 201
Db 421 GAGCGCGTGTCTTGTGACCTCATGACATGAGTGTGAGCATTAAGCAAGGACAAGGCG 480
QY 202 AlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 221
Db 481 GCTATCTCAGCAACACTGTGTAGACCAAGCGGAACACGGGACCAATCGCGCAGAGAG 540
QY 222 GluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnIle 241
Db 541 GAGCTGGCGCAGGCGCGGCTGTGTACTGAACTGCGACGATTTGACATTTGGAGACAGATC 600
QY 242 GlyGluGluGluPheGlyAlaValLeuGlnGlnGlyGlyTrpLeuGlnGlyValAlaVal 261
Db 601 GGAGAGGAGAGTTTGGAGCTGTCTCGAGGTGAGTACTGGGGCAAAAGGTGGCGCTG 660
QY 262 LysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMetThr 281
Db 661 AAGATATCAAGTGTGATGTACAGCCAGGCTTCCGAGAGAGAGCGCGCTCATGACG 720
QY 282 LysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeuTyr 301

```

Db      721 AAGATGCAACGAGAACTGTCGCTCCTCTGCGCTGATCCTGCACAGAGGCTGTAC 780
Qy      302 ILeValMetGluHisValSerLysGlyAsnLeuValAsnPhenLeuArgThrArgGlyArg 321
Db      781 ATGTGATGAGGACGTCAGACCAAGGGCAACCTGTGTGAATTTCTGCGAGCCCGGGGTGCA 840
Qy      322 AlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlnGlyMetGlu 341
Db      841 GCCCTCTGTAACACCGCTCAGCTCTCTCAAGTTTCTGTGCACTGGGCGGAGGCATGAG 900
Qy      342 TyrLeuGlnSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuValSer 361
Db      901 TACCTGGAGAGCAAGACTTGTGCACCGGCACTGGCCGCCCAACATCTGTGCTCA 960
Qy      362 GluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGlnGlyGlyLeu 381
Db      961 GAGGACCTGTTGGCCCAAGGTCAGAGCTTGGCTGCGCAAGCCGAGCGGAAGGGCTA 1020
Qy      382 AspSerSerArgLeuProValLysTrpThrAlaProGlnAlaLeuLysHisGlyLysPhe 401
Db      1021 GACTCAAGCCGCTGCTCCGTCMACTGACGCGCCCGAGGCTCTCAACACGGGAAGTTC 1080
Qy      402 ThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyrGly 421
Db      1081 ACCAGCAAGTCGATGTCGTGAGTTTGGGGTGTCTGGAGGTCCTCTCATATGGA 1140
Qy      422 ArgAlaProTyrProLysMetSerLeuLysGlyValSerGlnAlaGlnGlyTyr 441
Db      1141 CGGGCTCTGTAACCTAAAGTCACTGAAGAGTGTGCGAGGCGCGGGAAGGGGTAC 1200
Qy      442 ArgMetGluProProGlnGlyCysProGlyProValHisValLeuMetSerSerCysTrp 461
Db      1201 CGCATGTAACCCCGGAGGGCTGTCCAGGCGCCGCTGCACGTCCTCATGACAGCTGCTGG 1260
Qy      462 GluAlaGluProAlaArgArgProProPheArgLysLeuAlaGlnLysLeuAlaArgGlu 481
Db      1261 GAGGAGAGCCCGCCCGCGCCGACACCTTCGCAAACTGGCCGAGAAAGTCGGCCGGAG 1320
Qy      482 LeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThrSer 501
Db      1321 CTACGCAAGTCAGAGTGTGCTCCGCTCTGTCAGGGCAGGACCGCGCTCCACCTCG 1380
Qy      502 ProArgSerGlnGluPro 507
Db      1381 CCCGAAAGCCAGAGGCC 1398

RESULT 9
US-09-741-154-1
; Sequence 1, Application US/09741154
; Patent No. 6437110
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, ELLEN M. et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01061
; CURRENT APPLICATION NUMBER: US/09/741,154
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1713
; TYPE: DNA
; ORGANISM: Human
US-09-741-154-1

Alignment Scores:
Pred. No.: 2,16e-192 Length: 1713
Score: 2020.50 Matches: 410
Percent Similarity: 84.888 Conservative: 11
Best Local Similarity: 82.668 Mismatches: 48
Query Match: 75.654 Indels: 29
DB: 4 Gaps: 6

```

```

US-09-977-260-2 (1-507) x US-09-741-154-1 (1-1713)
Qy      23 ProArgValSerPro-----ArgPheLeuArgAlaTrp 33
Db      30 CCCCGAGACAGCGCGCGGTGCGAGAACACACAGCTGCAGGGGCTGAGAACCA 89
Qy      34 HisProProProValSerAlaArgMetProThrArgArgTrpAlaProGlyThrGlnCys 53
Db      90 CACCCCTCTCC-----TCCCCCAGAGAACTTGAGACCGGGGACCTCG--- 134
Qy      54 IleThrLysCysGlnHisThrArgProLysProGlyLysLeuAlaPheArgLysGlyAsp 73
Db      135 GTTGGCGA-TGC-----TGGCGGACCGCGG 160
Qy      74 ValValThrIleLeuGlnAlaCysGluAsnLysSerTrpTyrArgValLysHisThr 93
Db      161 TTGCTAGGACACATCGAGAGAGAGAGAGAAAGTTTATGAGGTGCTGACAGCATCTT 220
Qy      94 SerGlyGlnGluGlyLeuAlaAlaGlyAlaLeuArgGlnArgGlnAla---LeuSer 112
Db      221 AATG-TCCCTCTCATGCAAGGACATTTCCCTCTGAAACGAAAGAGACGACCTCGGC 279
Qy      113 AlaAspProLysLeu---SerLeuMetProTrpPheHisGlyLysLysSerGlnGln 131
Db      280 GAGGAGCCCGTGGTCAGACACAGCTTGTGGTGTCCAGCGGAAGATCTTGGCCAGAG 339
Qy      132 AlaValGlnGlnLeuGlnProProGluAspGlyLeuPheLeuValArgGlnSerAlaArg 151
Db      340 GCTGTCAGACAGTGCAGCGCTCCCGAGAGATGGGCTGTCTGTGTCGCGAGTCCGCGCC 399
Qy      152 HisProGlyAspTyrValLeuCysValSerPheGlyArgAspValIleHisTyrArgVal 171
Db      400 CACCCCGGCACATACGTCCTGTGAGCTTGGCGCGACGCTCATCCACTCCCGT 459
Qy      172 LeuHisArgAspGlyHisLeuThrIleAspGlnAlaValPhePheCysLeuMetAsp 191
Db      460 CTCGACCGGACGCGCCTCACAATGATGAGGCGGTGTTCTTGTCAACCTCATAGGAC 519
Qy      192 MetValGlnHisTyrSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLys 211
Db      520 ATGTGGAGCATTAACAGCAAGAGAGGGCGCTATCTGCCACCAAGCTGTGGAGAACCAAG 579
Qy      212 ArgLysHisGlyThrLysSerAlaGlnGluGlnLeuAlaArgAlaGlyTrpLeuAsn 231
Db      580 CGGAAACAGCGGACCAAGTGGCGGAGGAGAGACTGGCGGCGGTGTACTAGAAC 639
Qy      232 LeuGlnHisLeuThrLeuGlnAlaGlnIleGlyGlnGluPheGlyAlaValLeuGln 251
Db      640 CTCGACATTTTGACATTGGGAGCAGATCGGAGGAGAGAGTTTGAGCTGCTGCAG 699
Qy      252 GlyLysLysLeuGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGln 271
Db      700 GGTGAGTACTGGGGCAAAAGTGGCGCGTGAAGAAATATCAAGTGTATGTGACAGCCAG 759
Qy      272 AlaPheLeuAspGluThrAlaValMetThrLysMetGlnHisGlnAsnLeuValArgLeu 291
Db      760 GCCTTCCTGACAGAGCGCGCTCATGAGAGATCAACACAGAACCTGTGGCTCTTC 819
Qy      292 LeuGlnValIleLeuHisGlnGlyLeuTyrIleValMetGlnHisValSerLysGlyAsn 311
Db      820 CTGGGCGTATCTTGACCAAGGGGCTGTACATTGTATGAGAGCACTGAGCAAGGGCAAC 879
Qy      312 LeuValAsnPhenLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuGln 331
Db      880 CTGTGTAACCTTCTGGGAGACCGGGGTGACGCTCTGTGAACACCGCTAGCTCTGCAG 939
Qy      332 PheSerLeuHisValAlaGlnGlyMetGlyTyrLeuGlnSerLysLysLeuValHisArg 351
Db      940 TTTTCTCTCAGCTGCGCAGAGGATGAGTACTGTGAGAGCAAGACTTGTGCACCGC 999
Qy      352 AspLeuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaLysValSerAspPhe 371

```

Db 1000 GACCTGGCCGCCGCAACATCTGTCAGAGAGACCTGTCGCGCAAGTCAAGCAGCTTT 1059
QY 372 G1YleuAlaYsAlaGluArgLySgLYleuAspSerSerArgLeuProValLysTrpThr 391
Db 1060 GGCCTGGCCAAAGCCGAGCGGAGGCTAGACATCAAGCCGCTGCGTCAAGTGGACG 1119
QY 392 AlAProGluAlaLeuYsHISgLYLysPheThrSerLYsSerAspValTIPSerPheLY 411
Db 1120 GCGCCGAGGCTCTCAAAACGAGAGTTCACACAGCAAGTCGAGTCTGAGATTGTTGG 1179
QY 412 ValLeuLeuTrpGluValPheSerTrpGlyArgAlaProLYProLYMetSerLeuLYs 431
Db 1180 GTGCTGCTGTGGAGGTCCTCTCATATGACGGGCTCCGTACCTCAAAATGTCACCTGAA 1239
QY 432 GluValSerGluAlaValAlaGluLYSgLYTrpArgMetGluProProGluGlyCysProGly 451
Db 1240 GAGGTGTCGAGGCGCTGGAAGAGGGTACCGCATGCAACCCCGGAGGCTGTCCAGGC 1299
QY 452 ProValHISValLeuMetSerSerCysTrpGluAlaProAlaArgArgProPhe 471
Db 1300 CCCGTCACCTCTCATATGACAGCTGCTGGAGGACAGCCCGCCGCGCCACCTTC 1359
QY 472 ArgLYsLeuAlaGluLYsLeuAlaArgGluLeuArgSerAlaGlyAlaProAlaSerVal 491
Db 1360 CGCAAACTGCGGAGAGCTGGCCGGAGCTACGAGTCAGAGTGGCCCGCAGCTCCGTC 1419
QY 492 SerGlyAspAlaAspGlySerThrSerProArgSerGlnGluPro 507
Db 1420 TCAGGGCAGAGCGCGAGGCTCCACCTGCGCCCGCAAGCCAGAGCCC 1467

RESULT 10
US-09-741-154-3
Sequence 3, Application US/09741154
Patent No. 6437110

GENERAL INFORMATION:
APPLICANT: BEASLEY, Ellen M. et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEROP
FILE REFERENCE: C1001061
CURRENT APPLICATION NUMBER: US/09/741,154
CURRENT FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 3
LENGTH: 16389
TYPE: DNA
ORGANISM: Human
US-09-741-154-3

Alignment Scores:
Pred. No.: 1,03e-135 Length: 16389
Score: 1469.50 Matches: 501
Percent Similarity: 21.59% Conservatave: 0
Best Local Similarity: 21.59% Mismatches: 6
Query Match: 55.02% Indels: 1819
DB: 4 Gaps: 12

US-09-977-260-2 (1-507) x US-09-741-154-3 (1-16389)

QY 1 MetaLagIArgGlySerLeuValSerTrpArgAlaPheHISgLYCysAspSerAlaGlu 20
Db 7229 ATGGGSGGCGAGGCTCTGTTCTCTGGCGGCGATTTCAGCGCTGTGATTCTGCTAG 7288
QY 21 G1U----- 21
Db 7289 GAACTTCCCGGGTAAGATCACTTCCACAGGGGCTTGGGAGCCAGTTCTGGTCCC 7348
QY 21 ----- 21
Db 7349 ATCCAGATGCTCTGGGAGGCTCTCTGCTGCCCTCCAGCAGTCCGCCACCTGACCT 7408
QY 21 ----- 21

Db 7409 GGGCTGAAGGTGAGAGACCTGAGCCGTGGGTCCACAGCCAGCTCCACCTGAGACTCTGT 7468
QY 22 ---LeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
Db 7469 CCCCTCCAGAGTG-AGCCCCGCTCTCCAGAGCTGGAGACCCCTCCGCTCCAGCC 7527
QY 41 ArgMetProThr----- 44
Db 7528 AGGATGCCAACGGTGAAGTGTGTCTAGCCTGCTCTCTCTCCCGGGGCTCTTCTTAAC 7587
QY 44 ----- 44
Db 7588 TCCGTCACACCCCTGAGCCCTCTCCGAAAGGGCTGACTCTCTGCCCCCTACCCCTTC 7647
QY 44 ----- 44
Db 7648 TGCTGGCTCGGCTGGCTTTCACATCATCTTGGCTCTCTGACCTTCTGCCAGGCC 7707
QY 44 ----- 44
Db 7708 GCTCTCTTGTACTACATGATTTCTCTGAGCACCACCTCCCTTTCTTTCCCC 7767
QY 44 ----- 44
Db 7768 GATGTCGCGTGTCTTCTCTGCTCCGCTCTTTCACATCTATACCTTCTTCCC 7827
QY 44 ----- 44
Db 7828 TCTGTCTCTTTCCTCCCTCCGCTCTGCTCCACATCTCCGTATGTCCTCTT 7887
QY 45 ----- 44
Db 7888 GCCCTCTCCGCGCCCTCTGACAGAGCGCTGGGCCCGCGGACCCAGTATACACCA 7947
QY 56 sCysGluHISThrArgProLYsProGlyGluLeuAlaPheArgLYsGlyAspValAla 76
Db 7948 ATGGAGAGACACCGCCCGCAAGCCAGGAGACTGCTCTCCGAAAGGGCGAGCGTCCAC 8007
QY 76 rIleLeuGluAlaCysGlu----- 82
Db 8008 CATCTGAGGCTGCGCA-GGTGAGAGTGGCGCGGGTGTGGTCTTGGGCTGGGG 8066
QY 83 -----As 83
Db 8067 GCTCCACAGACACACCCACACCCCTCACTAACCCCTGCTCTCTCCCTGGCAGAA 8126
QY nLYsSerTrpTrpArgValLYsHISThrSerGlyGlnGluLYleuLeuAlaGlu 103
Db 8127 CAAGAGCTGTACCGCGTCACAGCACACACCATGAGAGAGGGCTGTGGACGCTGG 8186
QY 103 yAlaLeuArgGluArgGluAlaLeuSerAlaAspProLYsLeuSerLeuMetPro----- 121
Db 8187 GGGCTGGGAGGCGGGAGGCCCTCTCCGAGACCCCAAGCTCAGCTATGCGGTGAGT 8246
QY 121 ----- 121
Db 8247 GGGCAGACAGAGGGCTGGGCTAGGGAGACAGTACCCCTCCACAGCCAGCTCT 8306
QY 122 -----TrpHisGlyLYsIleSerGlyGlnGluAlaVal 133
Db 8307 GACCACCCCTTCGCTGGCGCCAGAGTGTTCACAGGGAATCTCGGGCCAGAGGCTGT 8366
QY 133 IGLGlnLeuGlnProProGluAspGlyLeuPheLeuValArgGluSerAlaArgHisArg 153
Db 8367 CCAGCAGCTGACCTCCCGAGATGGGCTTCTCTGTTGGGAGAGTCCGCGCCACACC 8426
QY 153 oGlyAspTrpValLeuCysValSerPheGlyArgAspValIleHisTrpArgValLeu 173
Db 8427 CGGCGACTACGCTCTGTGCTGAGCTTTGGCGGAGACTATCCACTACCGCGTGTCCA 8486
QY 173 sArgAspGlyHISLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspMetVal 193

Db 8487 CCGGAGGCGGACCTCACAATCGATGAGCGCTTCTTCTGCAACCTCATGACATGCT 8546
QY 193 Lglu----- 194
Db 8547 GGAGGTGCTGCCACCCAGAGGCCCCGCTCCCTGCAGTGGGGAGCCAGGAGACTGC 8606
QY 194 ----- 194
Db 8607 GGAGACTCCCTACGTAGAGATGAGAGGGGGGCTACCTGCATCCCTTCCCAAGC 8666
QY 194 ----- 194
Db 8667 AGCTGGGAGACACATCCCTACCTCTGTGCTGATCACCAGGAGCCCTGAGCTCTGT 8726
QY 194 ----- 194
Db 8727 GCTTCCCAATGTGAGATGAGAGGGGTATTCCCAAGGCCCCCAAGCAACCCCAAGGCC 8786
QY 194 ----- 194
Db 8787 CCCATCCCATCTCAGATGGGAGAGAGAGACCCCAAGTAGATTCCTTAATGCA 8846
QY 194 ----- 194
Db 8847 GAAATAGGGGGGCTGATCCCTATGAGACTGCCCAATCTACAGTTGGCTGAGTCAGCCTG 8906
QY 194 ----- 194
Db 8907 TTCTACTCCAGGAGCTAGAGTACTCTCTGAGGAGACCCCAAGAACCCCAAGTTGAT 8966
QY 194 ----- 194
Db 8967 ACCAATCCAGAGACTCCCTCTCTTTGGCTTCCCTTCCCTCTCTATGAAACC 9026
QY 194 ----- 194
Db 9027 AGCCTTCTCTCTCTCCCGGTCCACCCACCCAGAGGCCAAGCCGGAAGCGAAC 9086
QY 195 ----- 195
Db 9087 AGGATCTGGGGTTCCTCCCTCGGGGCTGAGGCTGTCCACATCTCTGAGCA 9146
QY 195 ----- 195
QY 195 sTyrSerIysAspLysGlyAlaIleCysThrIysLeuValArgProLysArgLysHisG 215
Db 9147 TTACAGCAAGAGAGAGGGGCTATCTGCACCAAGCTGTGTAGACCAAGGGAACACGG 9206
QY 215 yThrIysSerIleGluGluLeuValIaArgAla----- 226
Db 9207 GACCAAGTGGCGCGAGAGAGAGCTGGCCAGGGG--TAGGGAGGCCCAAGAGGCAAGACC 9265
QY 226 ----- 226
Db 9266 CTTTCCCTAACCCAGTTAGCCAGTCCGGGAAGGAAGGCGCTGGGGCCCCGACCTGTG 9325
QY 226 ----- 226
Db 9326 AGGCCAGATCACCAGCCTGTCTCATGCCAGCTATGGGATAGAAGACCTGGGCTGCC 9385
QY 226 ----- 226
Db 9386 TCGGGGGTGGCCACAGACACAGATCAGTCTTTATTCATAGTCCCGCAGCTTGGCACT 9445
QY 226 ----- 226
Db 9446 CATCTGTACATCATCATCATCTGAGCCAGAAACCTGGCCTGAACCTCCACCCACCC 9505
QY 226 ----- 226
Db 9506 CATCTGTACAGATTCCTCTCTGGCCAGGTTCCCAAGACCTGTGAATCTTCTCCCTCC 9565
QY 226 ----- 226
Db 9566 CCTCCCCACAGCCAGACCTTCAGGCGCTCCCATTCAGACCGGCTCTTCCCAAGAGGCGAGT 9625

QY 226 ----- 226
Db 9626 CCCCTTCCAGCCAGAGCCTTGAGTGCAGCTCAAGTTATCTTCTATCATGAACCAT 9685
QY 226 ----- 226
Db 9686 GCCTCAGGAGAGAGAGGCCAGCAGCTCCACACCTTCCAGAACTGCTGCATGGCTG 9745
QY 226 ----- 226
Db 9746 AGCCACCTTTCAGCCTCATCTGCAAAAGCCCTTCTCCACACATCCACCCCTCCAGGT 9805
QY 226 ----- 226
Db 9806 CAACGTGATCATGTTTTCTTCATGCTCTGAGCCATTGTATATTCTGTCCCTTACCC 9865
QY 226 ----- 226
Db 9866 TGAATGCTTTTCCAGCCCTTAAGTAGTGAATCTCAATATTAAAGTCAAAAGCAAT 9925
QY 226 ----- 226
Db 9926 GCCTCCTTGACCCCATGCTGTCTCTCTTGAAGTATCTCTGTCTGGGGAAGTCTGTAC 9985
QY 226 ----- 226
Db 9986 CCATCAGGGATGGCATCCAGTGTAGAGACTAGCATCACCTGGCCATGATCTTAAAGC 10045
QY 226 ----- 226
Db 10046 CGTGCAATTTGAGGGGGACCTTCTGAGGAGAGAGAGCTGTGAGAACCGACAG 10105
QY 226 ----- 226
Db 10106 AGGCGAGTGGAGGGGACAACAGAGAGTCTTGGGAAGGGTCAGGGAGCTAGAGCC 10165
QY 226 ----- 226
Db 10166 TGAATATTTCTGAACTTTATAGGAGTCTCCATGTGTGTTTTTCTGTGTGTGTT 10225
QY 226 ----- 226
Db 10226 CTGGGGTGTACAGAGTAAAGAACTGTGTAGGAGAGGCCAGAGGGTATTTGGAGGAG 10285
QY 226 ----- 226
Db 10286 AAGAGAGCCAAAGAGAAAAGGCGAGCTTATCTGTGAGACTGAAGTCTTCTGGCAC 10345
QY 226 ----- 226
Db 10346 TACCCCTTGGGCTGACTGGAGATTCTGTGAGGGCAAGTGTGGGTGTGAGGCTCACG 10405
QY 226 ----- 226
Db 10406 CAGGAGGCCCCAGCTTGACTATTACACACACTATCCCTCCTGTAAAGCCCTTTTGT 10465
QY 226 ----- 226
Db 10466 TCAGTCAACCTGAGATAGTTCTGTGTTGCAACAACAATAATGTGTACAGAGGTGG 10525
QY 226 ----- 226
Db 10526 GTGTCTGTATATGTCCTAGGGGCAATGCCAGGCCAGCAGCCACTCAATGTGTATCA 10585
QY 226 ----- 226
Db 10586 CAGCTGCAGACTGTGAGACATCAGTAGTCTCTCACCAAGTAGAAGCCCAATGGGA 10645
QY 227 ----- 230
Db 10646 GGAATTTAGGGGGCCATTAACCCCTGTGACTGGGTCTTCACTGCACAGCGGGCTGGTACT 10705

QY 230 uAsnLeuGlnHisLeuThrLeuGlyAlaGlnIleGlyGlnIlePheGlyAla----- 248
 Db 10706 GAACTGCGAGCATTTGACATTTGGAGCACAGATCGAGAGGAGAGATTGGAGG-TGAGT 10764
 QY 248 ----- 248
 Db 10765 GCGGAGCTGGGTTGGGGGTGCTGAAGAGATGGAGTGCATATTGAGGCTTTGGAGTC 10824
 QY 248 ----- 248
 Db 10825 ACAGAGCTGATTTCAAAACCCATCTTAACCACTACTAGTTGTGCACTGAAGACGTGCA 10884
 QY 248 ----- 248
 Db 10885 AATCCTGTGAGCCTCAGTTTCCCCCTGTGAACAGGGGGTAATGGACGTGTCATGAGGA 10944
 QY 248 ----- 248
 Db 10945 GTAAGTGAACTCAGAGCGCTCAGTCAGACGTAGTACTTCTTCACACTGCCAGGCTC 11004
 QY 248 ----- 248
 Db 11005 TTGTATGGTCACTCCCACTGTTTGAATGTTGAACACACCATTCCTCAATTAATTATTC 11064
 QY 248 ----- 248
 Db 11065 TGAAGGCTTGGGGTTAGATCAGGGGTGCGAGACTCTGTGTAAGGCGCAGATTGTA 11124
 QY 248 ----- 248
 Db 11125 AATATTCTGCTTTGTGGGCCAGGTAGCTGCTGACCACTACTGAAATCCGCCCTGT 11184
 QY 248 ----- 248
 Db 11185 GAGTGAAGGAGCACAGACAATATGTAACATATGGGCTGGCCATGAAACTATATGAC 11244
 QY 248 ----- 248
 Db 11245 CGAATAATTGAATTTGTGTAATGCTCATGTCCTGTAACAGCACTCATCTTTGATT 11304
 QY 248 ----- 248
 Db 11305 TGTTTTACAAACATTTAAACATGCAAAAACCAAGCAGCGGTGCTCATGCTGGA 11364
 QY 248 ----- 248
 Db 11365 TCACAGCACTTTGTAGGTGCGGGGAGAGATCCCTTGAGCTCAGAGTTCAGGTGAG 11424
 QY 248 ----- 248
 Db 11425 CCTGGACAACGTAGCGAGACCCCATCTGTACAAAAAAGCCAGGACAGTGGAGCAT 11484
 QY 248 ----- 248
 Db 11485 GCCTGTACTCCAGTACTCAGAGAGCGGGGATGGAGAGATCCCTTGAACCCAGGAATTT 11544
 QY 248 ----- 248
 Db 11545 GAACCTGCAGTAGGTGTGATCAGCGCACCTGCACTCTACCTGGGTGAGAGATGAGACCC 11604
 QY 248 ----- 248
 Db 11605 TGTATCAATATTTTAAAAAATGTTGGCCGGGACAGTGGCTCATGGCTGTAAATCC 11664
 QY 248 ----- 248
 Db 11665 CAGCACTTGGGAGGCTGAGGGGGGAGATCAGAGGTATGAGATCGAGACATCTGCC 11724
 QY 248 ----- 248
 Db 11725 TAAATGTGTAACCCCATCTCTACTAAAAAATTAAGCCGGGCGTGG 11784
 QY 248 ----- 248

Db 11785 TGTGGGGCCCTGTAGTCCCACTACTTGGGAGGCTGAGGCGAGAGAAATGCCGTGACCTG 11844
 QY 248 ----- 248
 Db 11845 GGAGCAGAGCTTGCAGTGAAGAGATGGGCGCACTGCATCCAGGCTGGGCGACAGAG 11904
 QY 248 ----- 248
 Db 11905 CAAGACTCGCTCAAAATAAATTAATTAATTAATAAATAAATAAATGCAAAAAACATTTT 11964
 QY 248 ----- 248
 Db 11965 GAGCTGAAGGGTCTGCAAAAAAGAGTGAATGATTTGGCCCCGAGGGGGTGGGT 12024
 QY 248 ----- 248
 Db 12025 TCACAGACCCTGCTAGATGTTCTCTTTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 12084
 QY 248 ----- 248
 Db 12085 GAGATGAGTTCACTCTGTACCCAGGCTGAGTGCATGGCACAAATCTTGGCTCACTG 12144
 QY 248 ----- 248
 Db 12145 CAACCTTGCCCTCCCGGGTTCAATGATTTCTCTGCTCAGGCTCCGAGTAGTGGAT 12204
 QY 248 ----- 248
 Db 12205 TACAGCATGTGCCACACACACTGGCTAGTTTGTATTGTAATTAAGAGACAGGTTTCTC 12264
 QY 248 ----- 248
 Db 12265 CATGTTAGTGGGCTGGTCTTGAATCTCTGACCTAGTATCCGCTGCTCAGCTCC 12324
 QY 248 ----- 248
 Db 12325 CAAAGTGTGGGATTACAGCGGTGAGCCAGCCGCGCAGACGACAGATGTTCTGTATCTT 12384
 QY 248 ----- 248
 Db 12385 GGGTCTGAGCTTGCAGAGGGGCGAGGGGTGAGAGCATGGGGCTGTGTTCAAGTCT 12444
 QY 248 ----- 248
 Db 12445 GGGGACCCGCACTGTGGAATGTACATGGGACCCGCTGTGAGCGGGCTGTCCGTG 12504
 QY 248 ----- 248
 Db 12505 TCTCTGTCCCTGTTGTTGGGGTCCGTGATCTCACCCGCTGCCACCCACCCACCCCC 12564
 QY 249 -----ValLeuGlnGlyuTyrlLeuGlyGlnIleValAlaValIleLeuHisGlnIle 267
 Db 12565 AGCTGCTGCAAGGTGAGTACCTGGGCGAAAGGTGGCTGTGAAGATATCAAGTGA 12624
 QY 267 pValThrAlaGlnAlaPheLeuAspLutThrAlaValIleMetThr----- 281
 Db 12625 TGTGACACCCAGGCTCTCCGAGAGACCGGCTCATGAC-GTGAATCCCAAGGTGG 12683
 QY 281 ----- 281
 Db 12684 GCTGGGACCGTGGGACGGGGGGTCCACGCTGCCCCACGCCACCCACCGCCCC 12743
 QY 282 -----LysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnIle 300
 Db 12744 CAGGAGATGCAACACGAAACCTGGTGGTCTCTGGGCTGTGATCTTCACACAGGGGCT 12803
 QY 300 uTyrlIleValMetGlnHisValSerLys----- 309
 Db 12804 GTACATGTCTATGGAGCACGCTGAGCAAGGTGGGGGCCACAGCGGGGAGGGGCCCA 12863
 QY 310 -----GlyAsnLeuVal 313
 |||||

```
D 12864 CGCAGCGGAGAGCCCAACATCCGGGGCCATCTCCACCCCAACAGAGGCAACCTGCT 12923
Q 313 IASPhleuArghrArghrArglyAlaValAsnThraIagluLeuInpheSe 333
D 12924 GAACCTTCTGCGACCCCGGGGTGAGCCCTGTGTAACACCGCTGAGCTCTGCAAGTTTTC 12983
Q 333 rIeu----- 334
D 12984 TCTGTAACTGGGACTCTCAGGCTGCTGGGCACTGGGGTCTGTGGGGCGCAGATTCCAAG 13043
Q 334 ----- 334
D 13044 ATCCATCAAGGGGAAACTGAGGACAGGGGCGAGGAGGCTTTGTAGCTTTGGCCCAAG 13103
Q 334 ----- 334
D 13104 CACCCGAGCTTTCTGAGCCCTGAATGTGGGCACTGCGACCTGAGCCCACTGCCCC 13163
Q 335 -----HisValAlaGluGlyMetGluTyrLeuGluSerLysLysLeuValH 350
D 13164 CTACTACCCCGAGCGAGCTGGCCGAGGCGATGAGTACTGTGAGAGAGAAAGCTTTGTC 13223
Q 350 IASrAspLeuAlaAlaArghAsnIleLeuValSerGluAspLeuValAlaLysValSera 370
D 13224 ACCGCGACCTGGCCGCCGCAACATCCTGGTCTCAGAGGACCTGTGGCCAAAGTCAAGC 13283
Q 370 sPpHeGlyLeuAlaAlaLysAlaGluArghLysGlyLeuAspSerArghLeuProValLysT 390
D 13284 ACTTTGGCTGTGGCCAAAGCGAGCGGAGGAGGCTAGACTCAAGCGGCTGCCCTCAAGT 13343
Q 390 rPThraIaProGluAlaLeuLysHsGly----- 399
D 13344 GGAGGGGGCCGAGGCTCTCAACACGGGGTGAACCTGCCCTTCACATACCCCTGGGGCT 13403
Q 399 ----- 399
D 13404 TTGGGGTCCCCCAGCTCTGCTGATGACCTGGGTATGTCCCTTGGCTCTCTGAAACT 13463
Q 399 ----- 399
D 13464 CCAAGGCAATGGCTATGCTCTCCCAAGAACTTTGGCCCAATTTGTTCTCTGTAGCC 13523
Q 399 ----- 399
D 13524 CCCCCTGGGCTCAGTTTCTCCAGCTCTCAAAAAAGCGTGGCTCAGACTGAAGGAA 13583
Q 399 ----- 399
D 13584 GAATTAACAGAGCTTCTTATGCACTAAGGCTGAGCTAGACTCAACTGGGGGGCGGCTTG 13643
Q 399 ----- 399
D 13644 GGGGAGACTCAATTAGAGAGAGACAGAGAGTGAAGGTCACCCAGAGAGGCTTCTTG 13703
Q 399 ----- 399
D 13704 AGGAGAGGGGCTGGAGAGCAAAACCTAGAGGCTCCCTCTCAACCCCTCTGGGGCC 13763
Q 400 -----LysPheThrSerLysSerAspValTrrSerPheGlyAlaLeuLeuTrrPgluValP 418
D 13764 CACAGAGATTCAACAGACAGAGTGTGAGATTGTTGGGAGTGTGCTGTGGAGAGTCT 13823
Q 418 hEsErTyrGlyArghAlaProTyrProLysMetSer----- 429
D 13824 TCTCATATGAGCGGGCTCGGTACCTTAATATGCT-GAGGGGGGTCCAGAGAGCACTG 13882
Q 429 ----- 429
D 13883 GGTTCGGGGAGGTCCAGAGCGTGTGGCCCTGACCCCTGCGCAAGCTGCTGCCCACT 13942
Q 430 ---LeuLysGluValSerGluAlaValGluLysGlyTyrArghMetGluProGluGlyC 449
D 13943 CACTGAAGAAGGTGTGAGAGCGCTGTGAAGAAGGGTATACCGCATGGAACCCCGGAGGGCT 14002
```

```
Q 449 ySPGlyProValHisValLeuMetSerCysTrrPgluAlaGluProAlaArghArp 469
D 14003 GTCCAGGCCCCGTGACAGTCTCTATGAGACACTGCTGGAGAGGAGAGCCCGCCGCCGC 14062
Q 469 rOPrOPheArghLysLeuAlaGluLysLeuAlaArghGluLeuArghSerAlaGlyAlaProA 489
D 14063 CACCTTTCGCAAACTGGCCGAGAGCTGGCCCGGAGACTACGCAGTGCAAGTTCGCCAG 14122
Q 489 IASerValSerGlyGluAlaAspLysPglSerThrSerProArghSerGluPro 507
D 14123 CTTCTGCTCAGGCGAGAGCGCCGACGGCTCTCACTTGCCTCCCGCAAGCAGAGGCC 14178

RESULT 11
US-08-604-989A-8
: Sequence 8, Application US/08604989A
: Patent No. 5834208
: GENERAL INFORMATION:
: APPLICANT: Sakano, S.
: TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/604,989A
: FILING DATE: February 23, 1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Charles E. Miller
: REGISTRATION NUMBER: 24,576
: REFERENCE/DOCKET NUMBER: 1920-026
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-8864/9741
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 738 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA to mRNA
: ORIGINAL SOURCE:
: ORGANISM: human
: STRAIN: UT-7
: US-08-604-989A-8

Alignment Scores:
Pred. No.: 8, 81e-118 Length: 738
Score: 1269.00 Matches: 246
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 47.51% Indels: 0
Gaps: 0

US-09-977-260-2 (1-507) x US-08-604-989A-8 (1-738)

Q 233 GlnHisLeuThrLeuGlyAlaGluInIleGlyGluGlyGluPheGlyAlaValLeuGlnGly 252
D 1 CAGCATTTGACATTGGGAGACACAGATCCGAGAGGAGAGATTGGAGCTTCTCTCGAGGGT 60
Q 253 GlnTyrLeuGlyGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGlnAla 272
```

```

Db      61 GAGTACCTGGGCGCAAAAGGTGGCCGCTGAAGAAATATCAAGTGTGATGTGACAGCCGACGCC 120
Oy      273 PheLeuAspGluThrAlaValMetThrIysMetGlnHisGluAsnLeuValArgLeuLeu 292
Db      121 TTCCTGGAGACAGAGCGCCCTCATGACAGAGATGCAACACGACAACTGGTGGGTCTCCG 180
Oy      293 GlyValIleLeuHisGlnGlyLeuTyrIleValMetGlnHisValSerIysGlyAsnLeu 312
Db      181 GCGGTGATCTCTACACAGGAGGCTGTACATTTGATGAGACAGCTGAGCAAGGCAACCTG 240
Oy      313 ValAsnPheLeuArgThrArgIlyArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPhe 332
Db      241 GTGAACCTTCTCGGAGCCCGGGGTGAGCCCTGCGAAACCGCTCAGCTCTGAGATT 300
Oy      333 SerLeuHisValAlaGlnGlyMetGluTyrIleGlnIleSerIysLeuValHisArgAsp 352
Db      301 TCTGTGACAGTGGCGGAGGAGCATGTGATCTGTGAGAGACAAAGCTTGTGACCGCGAC 360
Oy      353 LeuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaLysValSerAspPheGly 372
Db      361 CTGGCCGCGCCGCAACATCTCTGCTCAGAGACCTGTGGCAAGGTGACGACTTTGGC 420
Oy      373 LeuAlaLysAlaGlnArgIlyGlyLeuAspSerSerArgLeuProValIlyStrPhrAla 392
Db      421 CTGGCCAAAGCCGAGCGGAGGGGCTAGACTCAAGCCGCTCCGCTCAAGTGGACGGCG 480
Oy      393 ProGluAlaLeuLysHisGlyLysPheThrSerIysSerAspValIlyTrpSerPheGlyVal 412
Db      481 CCCGAGGCTCTCAACACGGGAGAGTTCACACAGCAGTGGAGTGTGGAGTTTGGGGTG 540
Oy      413 LeuLeuTyrPgluValPheSerTyrGlyArgAlaProTyrProLysMetSerLeuLysGlu 432
Db      541 CTGCTCTGGAGGTCTTCTCATATGACAGCGGCTCCGTCAAAATGTCACATGAAGAG 600
Oy      433 ValSerGluAlaValGlnLysGlyTyrArgMetGlnProProGlnGlyCysProGlyPro 452
Db      601 GTGTGGAGCGCCGTGAGAGAGGGGTACCGCATGGAACCCCGGAGGGCTGTCCAGGCCCC 660
Oy      453 ValHisValLeuMetSerSerCysTyrPgluAlaGluProAlaArgProProPheArg 472
Db      661 GTGCACGTCTCATGACAGCTGTGGAGGAGGACGCGCCGCGCGCCACCTTCCGC 720
Oy      473 LysLeuAlaGluLysLeu 478
Db      721 AAACGTGCCGGAAGACTG 738

```

RESULT 12
 US-07-820-011A-1
 Sequence 1, Application US/07820011A
 Patent No. 5336615
 GENERAL INFORMATION:
 APPLICANT: Bell, Leonard
 APPLICANT: Madril, Joseph A.
 APPLICANT: Warren, Stephen L.
 APPLICANT: Luthringer, Daniel J.
 TITLE OF INVENTION: Genetically Engineered
 TITLE OF INVENTION: Endothelial Cells Exhibiting Enhanced
 TITLE OF INVENTION: Migration
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Maurice M. Klee
 STREET: 1951 Burr Street
 CITY: Fairfield
 STATE: Connecticut
 COUNTRY: USA
 ZIP: 06430
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 5.25 inch, 360 Kb storage
 COMPUTER: IBM PC XT
 OPERATING SYSTEM: PC-DOS/MS-DOS 2.10
 SOFTWARE: Displaywrite 3
 CURRENT APPLICATION DATA:

```

? APPLICATION NUMBER: US/07/820,011A
? FILING DATE: 19920106
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Klee, Maurice M.
? REGISTRATION NUMBER: 30,399
? REFERENCE/DOCKET NUMBER: LB-101
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (203) 255 1400
? TELEFAX: (203) 254 1101
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1602 base pairs
? TYPE: NUCLEIC ACID
? STRANDEDNESS: Double
? TOPOLOGY: Linear
? MOLECULE TYPE: cDNA to mRNA
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? ORIGINAL SOURCE:
? ORGANISM: Gallus, gallus
? PUBLICATION INFORMATION:
? AUTHORS: Takeya, Tatsuo
? TITLE: Structure and Sequence of the
? TITLE: Cellular Gene Homologous to the RSV src
? TITLE: Gene and the Mechanism for Generating the
? TITLE: Transforming Virus
? JOURNAL: Cell
? VOLUME: 32
? PAGES: 881-890
? DATE: March, 1983
?
? US-07-820-011A-1
?
? Alignment Scores:
? Pred. No.: 4,51e-64
? Score: 737.50
? Percent Similarity: 52.84%
? Best Local Similarity: 36.40%
? Query Match: 27.61%
? DB: 1
? Gaps: 15
?
? US-09-977-260-2 (1-507) x US-07-820-011A-1 (1-1602)
?
Oy      23 ProArgValSerProArgPheLeuArgAlaTrpHisProProProValSerAla----- 40
Db      137 CCCCGAGCGGCTCT-----TTGGGACCGTGGCCAGCCGCAAGCTCTTGGGGGCT 190
Oy      41 -----ArgMetProThrArgArg-----TriAlaProGly--- 50
Db      191 TCACACTTCGACACCGCTTACGTCGCCGACGCGCGGGGCACTGGCTGC-GGCGTC 249
Oy      51 ThrGlnCysIleThrLysCysGlnHisThrArgProLysProGlyGluLeuAlaPheArg 70
Db      250 ACCACTTTCGTGGCTCTCTACGACACGATGCCGAGTGAACGAGCACTTCCTTCAAG 309
Oy      71 LysGlyAspValValThrIleLeuGlnLalacysGluAsnLysSerTyrTyrValLys 90
Db      310 AAAGGAGAACCGCCGACAGATTGTCAACACAGGAAGT---GACTGTGGCTGGCTCAT 366
Oy      91 HisHisThrSerGlyGlnGlyLeuLeuAlaAlaGlyAlaLeuArgGluArgGluAla 110
Db      367 TCCCTCATTACAGAGACAGAGCGGCTACATCCCGATACATMTGCGCCCTCAGACATCC 426
Oy      111 LeuSerAlaAspProLysLeuSerLeuMetProThrPheHisGlyLysIleSerGlyGln 130
Db      427 ATCCAGCGCTGAA-----GAGTGTACTTGGGAAGATCATCTGCTGG 468
Oy      131 GluAlaValGlnGlnLeuGlnProProGluAsp-----GlyLeuPheLeuValArgGlu 148
Db      469 GAGTCCGAGCGGCTGCTCTCAACCCGCAAAACCCCGGGGGAACCTTCTTGGTCCGGAG 528
Oy      149 SerAlaArgHisProGlyAspTyrValLeuGlyValSer-----PheGly 163

```

```

Db      529 AGCGAGACGACAAAGGTGCTTATGCTTCCGTTTCTGACTTTGACAAAGCCCAAGGGG 568
Qy      164 Argspvalillehistrargvalleuhisargasp---glyhisleuthrilleaspglu 182
Db      589 CTCATGTGTAAGACATACAGATCCGACAGCTGACAGCGCGGCTTCTACATCAGCTCA 648
Qy      183 Alvalapherphesysanleumetaspmetvalgluhtystrlysaspplsglyala 202
Db      649 CGCACACAGTTCAGACGCTGACAGAGCTGTGGCCCTACACTCCAAACATGCTGATGCG 708
Qy      203 Ilecysnhrlyslau-----Valargprolysargylshisgltyrlysserlialglu 220
Db      709 TTGTGTCACCGCTGACCAAGCTGTGCGACGTCACCAAGCCCAAGCCAGGGA----- 762
Qy      221 GlugluleuAlaarglalegltprleuAasnleuGlnhtsleuthrleuGlyAlaGln 240
Db      763 -----CTCGCCCAAGGAGCGGTGGAAATCCCGGAGAGCGCTGGCGCGTGGAGGTGAG 816
Qy      241 IlleglygluglupheglialavalleuGlnGlylutyrluGly---GlnlyVal 259
Db      817 CTGGGGCAGGGCTGCTTGGAGAGGCTGTGATGGAGGAGCTGGAAGCGCAACAGCAGAGTG 876
Qy      260 Alvallylsasnleuys---CysaspvalthrAlaGlnAlaheleuaspluthrAla 278
Db      877 GCCATTAAGACTGTGAAGCCCGGCAACATGTCCCGGAGGCGCTTCGAGGAAGCCCAA 936
Qy      279 ValmethrlysmetGlnhtsGlnAasnleuValargleuGlnlyAlahtleuhtsGln 298
Db      937 GTGATGAAGAAGCTCCGGCATGAGAGCTGTTCAGCTGACAGCAGAGGTGTGGGAAGAG 996
Qy      299 GlyleuYrllleValmetGlnhtsValserlysglyAasnleuValasnpheluaYrthr 318
Db      997 CCCATCTACACTGCTACTGTACATGACAGGAGGAGCGCTCCGTGATTTCTGAAGGGA 1056
Qy      319 ArglyArglaleuvalasnthrAlaGlnleuGlnpheserleuhtsValAlaGln 338
Db      1057 GAATGGGCAAGTACCTCGCGCTGCCACAGCTGTGATGTGCTGTGCTGATGCTGCC 1116
Qy      339 GlymetGlnlyrleuGlnserlyslsleuValhtsArgsppluAlahtsArgspplu 358
Db      1117 GGCATGGGCTATGTGGAGAGATGAGTACGTGACCGAGAACCTGGGGGGGCGCAACATC 1176
Qy      359 leuvalserGlnaspLeuValAlahtsValserAspPhelyleuAlahtsAla----- 376
Db      1177 CTGTGGGGGGAACCTGTGTGCTCAAGGTGGCTGACTTTGGCTGGCAGCCCTCATCGAG 1236
Qy      377 -----GluArgylsGlyleuaspserserArgleuProvalysrphrAlaProglu 394
Db      1237 GACACGAGTACACAGCAGCGCAAGGTGCCAAGTTCCTCCATCAAGTGGACAGCCCGGAG 1296
Qy      395 AlaLeuhtshtsGlyshpserlysserAspvalrpserserphelyleuValleuLeu 414
Db      1297 GCAGCCCTATATGCGCGGTTCACACATCAAGTGGATGTGCGTCTGGCATCTGCTG 1356
Qy      415 TrpGluValpheserlyrGlyArgAlaProtyrProlysmetserleuGlyValser 434
Db      1357 ACTGACTGACCAAGGCGCGGTGCCATACCGAGGATGCTCAACAGGAGGTGCTG 1416
Qy      435 GluAlaValGluysGlyrArgmetGluProProGlnGlyysProGlyProvalhts 454
Db      1417 GACACAGTGAAGGGGCTACCGCATGCTCCGCGCGCCGAGCGCGAGTGGCTGCAT 1476
Qy      455 ValLeuhtserserCysTrpGluAlaGluProAlaArgProProphelarglyslu 474
Db      1477 GACCTCATGTGACAGTGTGGCGGAGGAGACCTGAGAGAGGGCGCACTTTTGTAGTACTG 1536
Qy      475 AlaGlnlysluAlahtsArgGluLeuArgserAlaGlyAlaProAlaSerGlyGln 494
Db      1537 CAGGCTTCTCTG-----GAG 1551
Qy      495 AspAlaAspGlyserThrserProArgserGln 505

```

```

Db      1552 GACTACTTACCTGACAGAGCCCGACGTACCAG 1584
RESULT 13
PCT-US93-00445-1
; Sequence 1, Application PC/TUS9300445
; GENERAL INFORMATION:
; APPLICANT: Bell, Leonard
; APPLICANT: Madril, Joseph A.
; APPLICANT: Warren, Stephen L.
; APPLICANT: Lubringer, Daniel J.
; TITLE OF INVENTION: Genetically Engineered
; TITLE OF INVENTION: Endothelial Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maurice M. Klee
; STREET: 1951 Burr Street
; CITY: Fairfield
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06430
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 760 KB storage
; COMPUTER: DELL 486/50
; OPERATING SYSTEM: DOS 5.0
; SOFTWARE: Displaywrite 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00445
; FILING DATE: 19930105
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/820,011
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Klee, Maurice M.
; REGISTRATION NUMBER: 30,399
; REFERENCE/DOCKET NUMBER: ALX-101PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 255 1400
; TELEFAX: (203) 254 1101
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1602 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Gallus, gallus
; PUBLICATION INFORMATION:
; AUTHORS: Takeya, Tatsuo
; AUTHORS: Hanafusa, Hidesaburo
; TITLE: Cellulargene Homologous of the
; TITLE: Gene and the Mechanism for Generating the
; TITLE: Transforming Virus
; JOURNAL: Cell
; VOLUME: 32
; PAGES: 881-890
; DATE: March, 1983
; PCT-US93-00445-1
Alignment Scores:
Pred. No.: 4,51e-64 length: 1602
Score: 737.50 Matches: 186
Percent Similarity: 52.848 Conservative: 84
Best Local Similarity: 36.408 Mismatches: 185
Query Match: 27,618 Indels: 57
DB: 5 Gaps: 15
US-09-977-260-2 (1-507) x PCT-US93-00445-1 (1-1602)

```

23 ProArgValSerProArgPheLeuAlaTrpHisProProProValSerAla----- 40
117 CTTGGGGGAGAACCTGCTGTGCAGGTGGCTGACTTTGGGCTGGCAGCCCTCATCGAG 1236
377 -----GluArgLysGlyLeuAspSerSerArgLeuProValLysTrpThrAlaProGlu 394
1237 GACACAGAGTACACAGCAGCGCAAGGTGCCAAGTTCCCAACCAAGTGGACAGCCCGCAG 1296
415 TrpGluValPheSerTrpGlyArgAlaProTrpProLysMetSerLeuLysGluValSer 434
1357 ACTGAGCTGACACCAAGGCGCGGTGGCCATACCAAGGATGTCACACAGGAGGTGCTG 1416
435 GluAlaValGluLysGlyTrpArgMetGluProProGluGlyCysProGlyProValHis 454
1417 GACCAAGTGGAGAGGCGTACCGCATGCCCTGCCCGCGAGTCCCGGAGTCCGCTCAT 1476
455 ValLeuMetSerSerCysTrpGluAlaGluProAlaArgArgProPheArgLysLeu 474
1477 GACCTCATGTGCCAGTGTGCGGAGGAGGACCTTAGAGAGCGGCCACTTTGAGTACCTG 1536
475 AlaGluLysLeuAlaArgGluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGln 494
1537 CAGCGCTTCTCTG-----CAG 1551
495 AspAlaAspGlySerThrSerProArgSerGln 505
1552 GACTACTTCACCTGACAGAGAGCCCAAGTACAG 1584

164 ArgAspValIleHisTrpArgValLeuHisArgAsp--GlyHisLeuThrIleAspGlu 182
589 CTCATGTGAGAGCACTACAAATCCGCAAGCTGGCAGAGCGCGCTTACATCACTCA 648
183 AlaValPhePheCysAsnLeuMetAspMetValGlnHisTrpSerLysAspLysGlyAla 202
649 CGCACACAGTTCAGAGCGCTGCAGACAGCTGTGGCTGCTACTCCAAACATGCGATGCG 708
203 IleCysThrLysLeu-----ValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
709 TTGGCCACCGCTGACCAACAGCTGCTCCGCTTCTGCACTTGGACACGCCAGAGGGA----- 762
221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
763 -----CTGCCCAAGAGCGCTGGGAATCCCGGAGGTGCTGGCTGGAGGTGAG 816
241 IleGlyGluGlyLysPheGlyAlaValLeuGlnGlyLysGlyLysGlyVal-----GlnLysVal 259
817 CTGGGGCAGGCGCTGTTGGAGAGGTGTGATGGGAGCTTGGAACGCCACACAGAGTG 876
260 AlaValLysAsnIleLys---CysAspValThrAlaGlnAlaPheLeuAspGluThrAla 278
877 GGCATTAAGAGCTGTAAGCCCGGCAACATGTCCCGGAGGCGCTTCTCCAGAGAGCCCA 936
279 ValMetThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGln 298
937 GTGATGAAGAGCTCCGCGATGACAGAGTGTTCAGCTGTGCGCAGTGGTGTGGAGAG 996
299 GlyLeuTrpIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThr 318
997 CCCATGTACATGTCACATGATACATGAGCAAGGGAGACCTTCGCGATTTCTCGAAGGA 1056
319 ArgGlyArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlu 338
1057 GAGATGGCGAAGTACCTCGGCTGCCAAGCTCGTATGATGCTGCATGATTTGCAATCC 1116
339 GlyMetGluTrpLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIle 358
1117 GGCATGCGCTTATGTGGAGAGATGACATGACCTGCACAGACCTTGGCGGCGCCCAACATC 1176
359 LeuValSerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAla----- 376

1177 CTTGGGGGAGAACCTGCTGTGCAGGTGGCTGACTTTGGGCTGGCAGCCCTCATCGAG 1236
377 -----GluArgLysGlyLeuAspSerSerArgLeuProValLysTrpThrAlaProGlu 394
1237 GACACAGAGTACACAGCAGCGCAAGGTGCCAAGTTCCCAACCAAGTGGACAGCCCGCAG 1296
415 TrpGluValPheSerTrpGlyArgAlaProTrpProLysMetSerLeuLysGluValSer 434
1357 ACTGAGCTGACACCAAGGCGCGGTGGCCATACCAAGGATGTCACACAGGAGGTGCTG 1416
435 GluAlaValGluLysGlyTrpArgMetGluProProGluGlyCysProGlyProValHis 454
1417 GACCAAGTGGAGAGGCGTACCGCATGCCCTGCCCGCGAGTCCCGGAGTCCGCTCAT 1476
455 ValLeuMetSerSerCysTrpGluAlaGluProAlaArgArgProPheArgLysLeu 474
1477 GACCTCATGTGCCAGTGTGCGGAGGAGGACCTTAGAGAGCGGCCACTTTGAGTACCTG 1536
475 AlaGluLysLeuAlaArgGluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGln 494
1537 CAGCGCTTCTCTG-----CAG 1551
495 AspAlaAspGlySerThrSerProArgSerGln 505
1552 GACTACTTCACCTGACAGAGAGCCCAAGTACAG 1584

RESULT 14
US-07-820-011A-3
Sequence 3, Application US/07820011A
Patent No. 5336515
GENERAL INFORMATION:
APPLICANT: Bell, Leonard
APPLICANT: Madril, Joseph A.
APPLICANT: Warren, Stephen L.
APPLICANT: Luthinger, Daniel J.
TITLE OF INVENTION: Genetically Engineered
TITLE OF INVENTION: Endothelial Cells Exhibiting Enhanced
TITLE OF INVENTION: and Plasmidogen Activator Activity
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 kb storage
COMPUTER: IBM PC XT
OPERATING SYSTEM: PC-DOS/MS-DOS 2.10
SOFTWARE: Displaywrite 3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/820,011A
FILING DATE: 19920106
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: LB-101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255 1400
TELEFAX: (203) 254 1101
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1611
TYPE: NUCLEIC ACID
STRANDEDNESS: Double

TOPOLOGY: Linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapien
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT: Chromosome 20
 PUBLICATION INFORMATION:
 AUTHORS: Anderson, Stephen K.
 AUTHORS: Gibbs, Carol P.
 AUTHORS: Tanaka, Akio
 AUTHORS: Kung, Hsing-Jien
 TITLE: Human Cellular src Gene:
 TITLE: Nucleotide Sequence and Derived Amino
 TITLE: Acid Sequence of the Region Coding for
 TITLE: the Carboxy-Terminal Two-Thirds of
 JOURNAL: Molecular and Cellular Biology
 VOLUME: 5
 ISSUE: 5
 PAGES: 1122-1129
 DATE: May, 1985
 PUBLICATION INFORMATION:
 AUTHORS: Tanaka, Akio
 AUTHORS: Gibbs, Carol P.
 AUTHORS: Arthur, Richard R.
 AUTHORS: Anderson, Stephen K.
 AUTHORS: Kung, Hsing-Jien
 AUTHORS: Fujita, Donald J.
 TITLE: DNA Sequence Encoding the
 TITLE: Amino-Terminal Region of the Human c-src
 TITLE: Protein: Implications of Sequence
 TITLE: Divergence among src-Type Kinase
 JOURNAL: Molecular and Cellular Biology
 VOLUME: 7
 ISSUE: 5
 PAGES: 1978-1983
 DATE: May, 1987
 US-07-820-011A-3
 Alignment Scores:
 Pred. No.: 5,11e-63 Length: 1611
 Score: 727.00 Matches: 188
 Percent Similarity: 49.01% Conservative: 83
 Best Local Similarity: 34.01% Mismatches: 185
 Query Match: 27.22% Indels: 98
 DB: 1 Gaps: 15
 US-09-977-260-2 (1-507) x US-07-820-011A-3 (1-1611)
 QY 24 ArgValSerProArgPhe-----LeuArgAlaIrrPhtsProPro----- 36
 DB 14 AGAGCAAGCCCAAGATGCCAGCCAGCGCGCCGCGAGAGCCCGCCGCGAGAACTGTC 73
 QY 37 -----Pro 37
 DB 74 ACAGCGCTGGCGGGGCGCTTCCCGCCCTCGCAGACCCCGCAGCAAGCAGCCTGGCGCG 133
 QY 38 ValSerAlaArgMetProThrArgArgTTPAlaPro----- 49
 DB 134 ACAGCGCAAGCGCGCGCCAGCGCGGC-TCGCGCCCGCGCGCGCCCGCAGGCCAAGCTGTC 192
 QY 50 -----GlyThrGlnCysIleThrLysCysGluHisThrArgProLysProGly 65
 DB 193 GAGAGCTTCAACTCTCGGAGACCGCTCACCTCCCGCAGAGGGCGCGCGCGCTGGCGCGGT 252
 QY 66 -----GluLeuAla 68
 DB 253 GGAGTGACACACTTGTGGCCCTGTATGACTATGACTATAGAGCAGAGACAGACCTGTCC 312

QY 69 PheArgLysGlyAspValValThrIleLeuGlnAlaGlyGlnAsnLysSerTyrArg 88
 DB 313 TTCAGAAAGCGAGCGCGCTCCAGATTGTCAACAAACAGAGGGA---GACTGGTGGCTG 369
 QY 89 ValLysHisThrSerGlyGlnGluLysLeuAlaIleGlyAlaLeuArgGluArg 108
 DB 370 GCCCACTGCTGCTGACAGAGACAGACAGGCTACATCCCGCAGCACTAGCGGCCCTCC 429
 QY 109 GluAlaLeuSerAlaAspProLysLeuSerLeuMetProThrPheHisGlyLysIleSer 128
 DB 430 GACTCCATCCAGCTGAG-----GAGTGATTTTGGCAAGATCACC 471
 QY 129 GlyGlnGluAlaValGlnGlnLeuGlnProProGluAsp-----GlyLeuPheLeuVal 146
 DB 472 AGACGGAGTCAAGAGCGCTTACTGTCATGACAGAAACCGAGAGGAGACTTCCTGCTG 531
 QY 147 ArgGluSerAlaArgHisProGlyAspTyrValLeuLysValSer----- 161
 DB 532 CGAAGAGTGAAGACACGAAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 591
 QY 162 PheGlyArgAspValIleHisTyrArgValLeuHisArgAsp---GlyHisLeuThrIle 180
 DB 592 AAGGGCTCAAGCTGAAGACACTAGATCGGAGAGCTGAGCGCGCGCTCTACATC 651
 QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnIleTyrSerLysAspLys 200
 DB 652 ACCTCCCGCACCCAGTTCAACACAGCTGCACAGACAGCTGAGTGGCTTCACTCAACACAG 711
 QY 201 GlyAlaIleCysThrLysLeu-----ValArgProLysArgLysHisGlyThrLysSer 218
 DB 712 GATGGCTGTGACCGGCTGACACCGCTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 771
 QY 219 AlaGluGlnGluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGlnHisLeuThrLeuGly 238
 DB 772 -----CTGGCAAGAGATGCTGCGGAGATCCCTCGGAGATCGCTGCGCTGCGAG 819
 QY 239 AlaGlnIleGlyGlnGluPheGlyAlaValLeuGlnGlyGlyTyrLeuGly---Gln 257
 DB 820 GTCAGCTGGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 879
 QY 258 LysValAlaValLysAsnIleLys---CysAspValThrAlaGlnIlePheLeuAspGlu 276
 DB 880 AGGCTGCGCATCAAAACCTGAAAGCTGGCAGCATGTCTCCAGAGGCTTCTCGCAGAG 939
 QY 277 ThrAlaValMetThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyAlaIleLeu 296
 DB 940 GCCCAGCTGATGAAGAAGCTGAGAGCAATGAGAACTGTGCTGCTGCTGCTGCTGCTGCT 999
 QY 297 HisGlnGlyLeuTyrIleValMetGluHisValSerLysGlyAsnLeuValAsnPheLeu 316
 DB 1000 GAGGAGCCCATTTACATCGTACAGGAGTACATGACAGAGGAGTGTGCTGCTGCTGCT 1059
 QY 317 ArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisVal 336
 DB 1060 AAGGGGGAACAGAGCAAGTACCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1119
 QY 337 AlaGlnGlyMetGluTyrLeuGlnGlnLysLysLeuValHisArgAspLeuAlaIleArg 356
 DB 1120 GCCTCAGGATGAGCGTACGTGAGCGAGTACGACTACGCTCACCGGAGCTTGTGTCAGCC 1179
 QY 357 AsnIleLeuValSerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAla 376
 DB 1180 AACATCTGTGTGGAGAACTGCTGCTCAAAAGTGGCGCACTTGTGCTGCTGCTGCTGCT 1239
 QY 377 -----GluArgLysGlyLeuAspSerArgLeuProValLysTyrThrPhe 392
 DB 1240 ATTGAAGACATGAGTACACGGCGCGCAAGGTGCAAAATTCCTCAATCAAGTGGAGCGGT 1299
 QY 393 ProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspValTyrPheSerPheGlyVal 412
 DB 1300 CCAGAAAGCTGCCCTGATGCGCGCTTACACATCAAGTGGAGAGTGTGCTGCTGCGGATC 1359
 QY 413 LeuLeuTyrGluValPheSerTyrGlyArgAlaProTyrProLysMetSerLeuLysGlu 432

DB 1360 CTGCTGACTGACTGCACACAAAGGACGGGTGCCCTACCTCGGATGTGTGAACGGCGAG 1419
QY 433 ValSerGluAlaValGlnLysGlyTyrArgMetGluProGluGlyCysProGlyPro 452
DB 1420 GGGCTGGACCAAGGAGGGGGGTACCGGATGCCCTGCCCGGAGGTGTCCCGACATCC 1479
QY 453 ValHisValLeuMetSerCysTyrPgluAlaGluProAlaArgArgProProPheArg 472
DB 1480 CTGCACGACCTCATGTGCCAGTGTGGCGGAGAGACCTGAGAGAGCGGCCACCTTCGAG 1539
QY 473 LysLeuAlaGluLysLeuAlaArgLysLeuArgSerAlaGlyAlaProAlaSerValSer 492
DB 1540 TACCTGCAGGCGCTCTCTG----- 1557
QY 493 GlyGlnAspAlaAspGlySerThrSerProArgSerGln 505
DB 1558 ---GAGGACTACTTACGTCACGACGAGCCCGACGTACGAG 1593
RESULT 15
PCT-US93-00445-3
Sequence 3, Application PC/TUS9300445
GENERAL INFORMATION:
APPLICANT: Bell, Leonard
APPLICANT: Madril, Joseph A.
APPLICANT: Warren, Stephen L.
APPLICANT: Luthringer, Daniel J.
TITLE OF INVENTION: Genetically Engineered
TITLE OF INVENTION: Endothelial Cells
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 760 Kb storage
COMPUTER: DELL 486/50
OPERATING SYSTEM: DOS 5.0
SOFTWARE: Displaywrite 3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00445
FILING DATE: 19930105
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/820, 011
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-101PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255 1400
TELEFAX: (203) 254 1101
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1611
TYPE: NUCLEIC ACID
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapien
POSITION IN GENOME:
CHROMOSOME/SEGMENT: Chromosome 20
PUBLICATION INFORMATION:
AUTHORS: Anderson, Stephen K.
AUTHORS: Gibbs, Carol P.
AUTHORS: Tanaka, Akio

AUTHORS: Kung, Hsing-Jien
AUTHORS: Fujita, Donald J.
TITLE: Human Cellular src Gene:
TITLE: Nucleotide Sequence and Derived Amino
TITLE: Acid Sequence of the Region Coding for
TITLE: the Carboxy-Terminal Two-Thirds of
TITLE: pp60c-src
JOURNAL: Molecular and Cellular Biology
VOLUME: 5
ISSUE: 5
PAGES: 1122-1129
DATE: May, 1985
PUBLICATION INFORMATION:
AUTHORS: Tanaka, Akio
AUTHORS: Gibbs, Carol P.
AUTHORS: Arthur, Richard R.
AUTHORS: Anderson, Stephen K.
AUTHORS: Kung, Hsing-Jien
AUTHORS: Fujita, Donald J.
TITLE: DNA Sequence Encoding the
TITLE: Amino-Terminal Region of the Human c-src
TITLE: Protein: Implications of Sequence
TITLE: Divergence among src-Type Kinase
TITLE: Oncogenes
JOURNAL: Molecular and Cellular Biology
VOLUME: 7
ISSUE: 5
PAGES: 1978-1983
DATE: May, 1987
PCT-US93-00445-3
Alignment Scores:
Pred. No.: 5 11e-63 Length: 1611
Score: 727.00 Matches: 188
Percent Similarity: 49.01% Conservative: 83
Best Local Similarity: 34.00% Mismatches: 185
Query Match: 27.22% Indels: 98
DB: 5 Gaps: 15
US-09-977-260-2 (1-507) x PCT-US93-00445-3 (1-1611)
QY 24 ArgValSerProArgPhe-----LeuArgAlaTrpHisProPro----- 36
DB 14 AGACCAAGCCCAAGAGGATGCCAGCGGGCGCCGACCTGGACCCCGGAGAACGTGC 73
QY 37 -----Pro 37
DB 74 ACGGCGTGGGGGGGGCTTCCCGCTCGACACCCCGACGACGACGCTGGCGG 133
QY 38 ValSerAlaArgMetProThrArgArgTrpAlaPro----- 49
DB 134 ACGGCCACCGGGGCCCGACGGCGGC--TTGCCCCCGCGCGCGCGGCCGACGCTGTC 192
QY 50 -----GlyThrGlnCysIleThrLysCysGlnHisThrArgProLysProGly 65
DB 193 GGAAGCTTCAACTCTCGGACACCGTCACCTCCCGGAGAGGGGGCGCGCGGT 252
QY 66 -----GluLeuAla 68
DB 253 GGAAGTACCACTTGTGGCCCTCTATGACTATGATGAGAGGAGAGAGACGCTGCC 312
QY 69 PheArgLysGlyAspValValThrIleLeuGlnAlaCysGlnAsnLysSerTyrPyrArg 88
DB 313 TTCAGAAAGGCGGCGGCTCCAGATGTGCACACACACAGAGGGA---GACTGTGGCTG 369
QY 89 ValLysHisHisThrSerGlyGlnGlyLeuLeuAlaAlaGlyAlaLeuArgGluArg 108
DB 370 GCCCACTCGCTCAGACAGACAGAGCGCTACATCCCGCACTACGTCGCGCTCC 429
QY 109 GluAlaLeuSerAlaAspProLysLeuSerLeuMetProThrPheHisGlyLysSer 128
DB 430 GACTCCATCCAGGCTGAG-----GAGTGTATTGTCAGAGATCAC 471

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2003, 15:38:14 ; Search time 448 Seconds
(without alignments)
3054.944 Million cell updates/sec

Title: US-09-977-260-2
Perfect score: 2671
Sequence: 1 MAGRSLVSMRAFHGCDSEAE.....PASVSGQADAGTSRSGEP 507

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 252756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_plus_p2n.model -DEV=tblp
-O=/cgn2_1/USPTO.spool/US09977260/rnatc.29072003_092740_14573/app.query.fasta_1.647
-DB=N.Geneseq.19jun03 -QFMT=fastap -SUFFIX=ing -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -STRT=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09977260.eccn_1.1.471.grunatc.29072003_092740_14573 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
N.Geneseq.19jun03:*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2671	100.0	1997	24	AA150631	Human matk gene-re
2	2668	99.9	1942	16	AA084888	DNA encoding cytop
3	2668	99.9	2000	16	AA1000516	Megakaryocyte kina
4	2591.5	97.0	1967	19	AAV44497	Human matk cDNA.
5	2020.5	75.6	1713	24	ABK8791	CDNA encoding huma
6	1540.5	57.7	1821	24	AB199884	Mouse ischaemic co
7	1469.5	55.0	16389	24	ABK8792	Gene encoding huma
8	1247.5	46.7	2187	21	AA229701	Wild-type human c-
9	1247.5	46.7	2187	22	AAH28358	Nucleotide sequenc
10	1247.5	46.7	2187	24	ABK84602	Human cDNA differe
11	1247.5	46.7	2187	24	ABL62918	Breast cancer rela
12	1247.5	46.7	2420	21	AA246489	PKA substrate, Csk
13	1247.5	46.7	2420	25	ACC50120	Breast cancer asso
14	1247.5	46.7	2420	22	AA193804	Human polynucleoti
15	1196.5	44.8	2442	24	ABK73326	DNA encoding human
16	994.5	37.2	2591	23	ABL11799	Drosophila melanog
17	832.5	31.2	14283	23	ABK81798	CDNA encoding tumo
18	749.5	28.1	2282	24	ABK5203	Drosophila melanog
19	744	27.9	4515	23	AB107083	PKA substrate, SRC
20	738.5	27.6	2032	21	AA464491	Chicken pp60 c-src
21	737.5	27.6	1602	14	AA046687	Wild-type chicken
22	735.5	27.5	1759	21	AA229700	Nucleotide sequenc
23	735.5	27.5	1759	22	AAH28357	Human pp60 c-src g
24	727	27.2	1611	14	AA046688	DNA encoding human
25	727	27.2	1611	24	ABK73325	Human src-c encodi
26	727	27.2	1611	25	ABK59382	Human signalling p
27	727	27.2	2129	25	ACA56854	Human cDNA differe
28	725.5	27.2	2015	24	ABK83939	Lung cancer relate
29	725.5	27.2	2015	24	ABK66673	Human cDNA differe
30	725	27.1	1926	24	ABK83940	Human cDNA differe
31	718.5	26.9	1466	25	ABK59389	Mouse src-c encodi
32	718	26.9	1466	25	ABK59752	Novel human coding
33	715.5	26.8	1911	24	ABK63704	Rat sequence diffe
34	710	26.6	4517	20	AA190200	Human yest encodin
35	710	26.6	4517	22	AAH28359	Nucleotide sequenc
36	710	26.6	4517	23	AAH24489	DNA encoding novel
37	710	26.6	4517	24	ABV94313	Breast carcinoma r
38	709.5	26.6	2298	24	ABK83935	Human cDNA differe
39	704	26.4	3327	23	AAK66986	DNA encoding novel
40	704	26.4	7487	23	AAK592457	Human gene express
41	702.5	26.3	4414	24	ABK35401	Human cDNA sequenc
42	701.5	26.3	1614	25	ABK74390	Human cDNA sequenc
43	701.5	26.3	5520	18	AAK61865	c-abl gene. Homo
44	701.5	26.3	5527	24	ABK97215	Gene #3713 used to
45	701	26.2	3258	24	ABQ74506	Mouse Fyn nucleoti

ALIGNMENTS

RESULT 1	AA150631	standard; DNA; 1997 BP.	
XX	AA150631:		
AC	02-JAN-2003	(first entry)	
XX			
DE	Human matk gene-related coding sequence.		
XX			
KW	Human; gene; ds; allergic disease; matk gene; atopic dermatitis.		
XX			
OS	Homo sapiens.		
XX			
PH	Key	Location/Qualifiers	
FT	CDS	266..1789	
FT		/*tag- a	

FT		/product= "Human matk-related protein"
XX		
PM	WO200275304-A1.	
PD	26-SEP-2002.	
XX		
PF	01-MAR-2002; 2002MO-JP01916.	
XX		
PR	21-MAR-2001; 2001JP-0081028.	
XX		
PA	(GENO-) GENOX RES INC.	
XX	(NIGE-) JAPAN GEN AGENCY NATION.	
PI	Sugita Y, Helshl M, Kagaya S, Gunji S, Tsujimoto G;	
XX		
DR	WPI: 2002-750572/81.	
XX	P-Psdb: AA015860.	
PT	Examining allergic diseases by changes in expression level of matk gene	
XX	in peripheral monocytes as indication, also applicable in screening	
PR	compounds for treating allergic diseases e.g. atopic dermatitis -	
XX		
PS	Disclosure; Page 48-52; 60pp; Japanese.	
XX		
CC	The invention comprises a method of examining allergic diseases, the	
XX	method comprises comprting the expression level of the matk gene in a	
CC	biological sample of a patient and a healthy individual. The method of	
CC	the invention is useful for examining/diagnosing allergic diseases -	
CC	particularly atopic dermatitis, the method of the invention is also	
CC	useful in screening candidate compounds for remedies. The present DNA	
CC	sequence encodes a human matk-related protein.	
SO		
Sequence	1997 BP; 372 A; 631 C; 676 G; 318 T; 0 other;	
Alignment Scores:		
Pred. No.:	1,22e-198	Length: 1997
Score:	2671.00	Matches: 507
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	24	Gaps: 0
US-09-977-260-2 (1-507) x AAL50631 (1-1997)		
OY	1 MetrlAGLVARGLISerLeuValSerTrPratgAlaPhenHISGLYCysASPSeRlaGlu 20	
Db	ATGGCGGGGGAGAGCCTCTGTTCCTGGCCGGCATTTCAACGCGTGAATTCTGCAGAG 3252	
OY	21 GlUdeProkrAvalSerProMrgrPhelauAgaIatRPHisPPROPtoVatSerAla 40	
Db	GAACCTGCCGGGTGAGCCCCCGCTTCTCCGACGCTGGCACCCCCCTCCGTTCAACC 3855	
OY	41 ArgMeerProThraRgtRtPralabrogILyThrInCylsIleThrLyScySGLNHISThr 60	
Db	AGGATGCGCAACGAGGGCGTGCGGCCCGGGCACCCAGTAGTATACACCAATGGGAGCACACC 4455	
OY	61 ArgProLyserogLyglUeuAlaPharGlySGlyAspValValThrIleUeuAla 80	
Db	CGGCCCAAGCCAGGGAGCGTGGCTTCGCCAACAGGGCGACGTGTACCATCTCGAGGCC 5050	
OY	81 CyegLIAsnLySerTPyrTyraRgtValLySHISHIHrSerGIyngIngLIyLeu 100	
Db	TGGGAGAACAAGCTGCTACCGCGCTCAAGCACACACCAAGTGGACAGAGGGGCTGCTG 5655	
OY	101 AlaAlaGLyAlaLeuArgLIuArgLIuAlaLeuSerAlaSProLySLeuSerLeuMet 120	
Db	GCAAGCTGGGGCGCTCGGGAGCGGAGGCGCTCTCCGACAGCCCAAGCTCAGGCTCATG 6255	
OY	121 ProTrphehIsGLySLeserGIyngIngLIuAlaValAIyngIngLnLeuInProProGlu 140	
Db	CCGTGTGTCCACGGGAAGATCTCGGGCCAAGGAGGCTGTCCAGCAAGCTGCAGCCCTCCGAG 6855	
OY	141 ASgGLIyLeuPhelauValArgLIuSerAlaArqHISPProGlyAspTYFValIleuCysVal 160	

Db	666	GATGGGCTGTTCTCTGGTGGGAGTCCGGCGCCACCCCGGCGACTACGTCTCTGGCTG	745
QY	161	SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle	180
Db	746	AGCTTTGGCCGCGAGCTCATCTACACTACCGCGCTGGCCACCGCGAGCGGCACCTCAATC	805
QY	181	AspGluAlaValPhePheCysAsnLeuMetSerPheTyrIguHisTyrSerLeuAspLys	200
Db	806	GATGGGCGCGTGTCTTCTGCACCTCTAGACATGGGAGCATTTACAGCAAGACACAG	865
QY	201	GlyAlaIleCysThrIleValLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu	220
Db	866	GAGCGTATCTGCACCAAGCTGTGTGAGACCAAGCGGAACACCGGAGCAAGTCCGGCAG	925
QY	221	GluGluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGlnHisLeuThrLeuGlyValAlaGln	240
Db	926	GAGGAGCTGGCCAGGCGGGCTGGTACTGAACTGCAGCATTTTGACATTGGAGCACAG	985
QY	241	IleGlyGluGlyGluPheGlyValAlaValLeuGlnGlyGlyTyrLeuGlyGlnLysValAla	260
Db	986	ATCGGAGAGGAGAGACTTTGGAGCTGTCTCTGCAGGGTGAAGTACCTGGGCAAAAGTGGCC	1045
QY	261	ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet	280
Db	1046	GTGAAGAATATCAAGTGTGATGTATACAGCCAGCGCTTCTCTGGACAGACCGCGCTATG	1105
QY	281	ThrLysMetGlnHisGluAsnLeuValArgPheLeuGlyValIleLeuHisGlnGlyLeu	300
Db	1106	ACGAAGATGCACACAGACAGACTGTGGCTGTCTCTGGGCGTGATCTTCACCAAGGGCTG	1165
QY	301	TyrIleValMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly	320
Db	1166	TACATTGTGCATGAGACAGCTGAGCAAGGCAACTGGTGAACCTTTGCGGACCCGGGGT	1225
QY	321	ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMet	340
Db	1226	CGAGCCCTCGTGAAACCCGCTACGCTCTCTGCAATTTCTCTGCACGTTGGCGAAGGCAATG	1285
QY	341	GluTyrLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal	360
Db	1286	GAGTACCTGGAGAGCAAGAACCTTGTGTGACCGCGACCTGGCCGCGCAATCATCTGGTTC	1345
QY	361	SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly	380
Db	1346	TCAGAGGACCTGGTGGCCAAAGTCAGCACATTGGCTGGCCAAAGCCGAGCAAGGAGCGG	1405
QY	381	LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys	400
Db	1406	CTAGACTACACCGCGCTGCCCTCAAGTGGAGCGGCGCCGAGGCTCTCAACACGGGAG	1465
QY	401	PheThrSerLysSerAspValTyrSerPheGlyValLeuLeuTyrPgluValPheSerTyr	420
Db	1466	TTACACACAGATGGAGTGTGTGGAGTTTGGGGTGTCTCTGGAGAGTCTTCTCATAT	1525
QY	421	GlyArgAlaProTyrTrpLysMetSerLeuLysGluValSerGluAlaValGluLysGly	440
Db	1526	GGAGCGGCTCGTACCTTAATAATGTCACTGAAGAGTGTGGAGGCGCTGTGAAGAGGG	1585
QY	441	TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerSerCys	460
Db	1586	TACCGCATGGAACCCCGGAGGCTGTCCAGGCCCGTCGACCTCTCATGACAGCTTC	1645
QY	461	TrpGluAlaGluProAlaArgArgProTyrPheArgLysLeuAlaGluLysLeuAlaArg	480
Db	1646	TGGGAGGCGAGCCCGCCCGCGCCACCTTCGCGAAACTGGCCGGAAGACTGGCCGG	1705
QY	481	GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr	500
Db	1706	GAGCTACGACTGAGGTCGCCCAAGCCTCTGCTCAAGGACGAGCGCGCAAGGCTCCACC	1765
QY	501	SerProArgSerGlnGluPro	507

Db		1766 TCGCCCAAGCCAGAGGCC	1786
RESULT 2			
AAQ84888			
ID	AAQ84888 standard; cDNA to mRNA;	1942 BP.	
XX			
AC	AAQ84888;		
XX			
DT	25-MAR-2003	(updated)	
DT	26-OCT-1995	(first entry)	
XX			
DE	DNA encoding cytoplasmic tyrosine kinase.		
XX			
KW	cytoplasmic; tyrosine kinase; blood; cell differentiation;		
KW	screening; anticancer agent; ds.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
CDS		208..1731	
FT	/tag= a		
FT	/product= cytoplasmic-tyrosine_kinase		
FT	/note= "see AAR7113"		
FT	349..540		
FT	/tag= b		
FT	/note= "encodes SH3 domain (see AAR71129)"		
FT	571..795		
FT	/tag= c		
FT	/note= "encodes SH2 domain (see AAR71130)"		
FT	904..1641		
FT	/tag= d		
FT	/note= "encodes tyrosine kinase domain (see AAR71131)"		
FT	misc_feature	331..1728	
FT	/tag= e		
FT	/note= "encodes N-terminally truncated form of the enzyme (see AAR71132)"		
XX			
PX	WO9506113-A1.		
PD	02-MAR-1995.		
XX			
PF	25-AUG-1994;	94WO-JP01411.	
XX			
PR	25-AUG-1993;	93JP-0210403.	
PR	29-MAR-1994;	94JP-0058553.	
PA	(ASAH) ASAMI KASEI KOYO KK.		
PI	Sakano S;		
DR	WP1; 1995-106842/14.		
DR	P-FSDB; AAR71129, AAR71130, AAR71131, AAR71132, AAR71133.		
PT	Cytoplasmic tyrosine kinase and antibody recognising it - for screening chemical substances for tyrosine kinase inhibitory or activating activity for use as cancer therapy		
PS	Claim 7; Page 49-50; 58pp; English.		
XX	This DNA encodes a cytoplasmic tyrosine kinase which has enhanced expression in connection with blood cell differentiation. It was isolated from the human um-7 blood cell line. The DNA sequences and antibodies raised against the enzyme, are useful for screening agents for inhibiting or activating activity on the tyrosine kinase, for use as anticancer agents.		
CC	(Updated on 25-MAR-2003 to correct PN field.)		
Seq	Sequence 1942 BP; 365 A; 615 C; 651 G; 311 T; 0 other:		
Alignment Scores:			
Pred. No. :	2.03e-198	Length:	1942
Score:	2668.00	Matches:	506

Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.80%	Mismatches:	0
Query Match:	99.89%	Indels:	0
DB:	16	Gaps:	0
US-09-977-260-2 (1-507) x AAG84888 (1-1942)			
QY	1	MetaIaGIyAArgGlySerIleuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu	20
Db	208	ATGGGGGGGGCGAGGCTCTCTGGTTCTCGGCCGGGCAATTTCACAGGCGGTGATTCCTGCTGAG	267
QY	21	GluteuProArgValSerProArgPheIleuArgAlaTrpHisProProArgAla	40
Db	268	GAAC TTCCTCCCGGGGTAGCCCTCCGCTTCCTCCGAGCTGCACCCCTCCCTCCGTACGCC	327
QY	41	ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr	60
Db	328	AGGAGTCCACAGAGCGCTGGGCCCCGGGACCCAGCTGATACCAAAATGCGAGCACACC	387
QY	61	ArgProLysProGluGluIleuAlaPheArgGlyGlyAspValValTrpIleIleuGluAla	80
Db	388	CGCCCCAACCCAGGGGAGCTGGCTTCCTCCGAAAGGGAGAGTGTGTACCAATCTCGGAGGCC	447
QY	81	CysGluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyIleGlnGlyLeuIleu	100
Db	448	TGGCCACCAAGAGCTGGTACCGCTCAAGCACCAACACAGTGGACAGGAGGGGCTGTG	507
QY	101	AlaAlaGlyAlaIleuArgGluArgGluAlaIleuSerAlaAspProLysIleuSerIleuMet	120
Db	508	GCAGCTGGGGGCTCGCGGAGCGGGAGGCCCTCTCGCAGACCCCAAGCTCAAGCTCCAG	567
QY	121	ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnIleuGlnInProProGlu	140
Db	568	CCGAGGTTCACAGGGAGATGTGGGGCCAGAGGCGGTGTCCAGAGCTGCAGCTCCGCGAG	627
QY	141	AspGlyLeuPheIleuValArgGluSerAlaArgHisProGlyAspTrpValIleuCysVal	160
Db	628	GATGGGCGTGTCTGTGTCGGGAGTCCGCGGCCACCCCGGAGCTACGTCTCTGTGCTGTG	687
QY	161	SerPheGlyArgAspValIleHisTrpArgValIleuHisArgAspGlyHisIleuThrIle	180
Db	688	AGCTTTGGCCCGGACGATATCCACTACCGCGTGTGCACCGGAGGCGCCACCTCACAAATC	747
QY	181	AspGluAlaValPhePheCysAsnIleuMetAspMetValGluHisTrpSerLysAspLys	200
Db	748	GATGAGGCGCGTGTCTCTGTGCACCTCATGCACTGAGCATGGAGCATTAAGCAAGCAAG	807
QY	201	GlyValAlaIleCysTrpThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu	220
Db	808	GGCGCTATCTCACCAACAGCTGTGTGAGCCAAAGCGGAACACGGGACCCAAAGTGGCGCGAG	867
QY	221	GluGluIleuValArgAlaGlyTrpPheIleuAsnIleuGlnHisLeuThrIleuGlyAlaGln	240
Db	868	GAGGAGCTGGCCAGCGCGGCTGTACTGAACCTGACAGCATTTGACATTTGGAGACAG	927
QY	241	IleGlyGlnGlyGluPheGlyAlaValIleuGlnGlyIleuTrpLeuGlyGlnLysValAla	260
Db	928	ATCGGAGGAGGAGGTTTGAGAGTCTCTGTGACGGGTGATCACTGGGCAAAAGGTGGCC	987
QY	261	ValLysAsnIleLysCysAspValTrpThrAlaGlnAlaPheIleuAspGluThrAlaValMet	280
Db	988	GTGAAGATATCAAGTGTGATGTGACACCCAGGCTTCCTGTGAGAGAGAGGCGCTCAG	104
QY	281	ThrLysMetGlnHisGluAsnLeuValArgLeuIleuGlyValIleuHisGlnGlyLeu	300
Db	1048	ACGAGAGATGCACACGAGAACCTGTGTGCTCTCGGCGCTGATCTGCACACAGGGCGGT	1107
QY	301	TyrTrpIleValMetGluHisValSerLysGlyAsnLeuValAsnPheIleuArgTrpArgGly	320
Db	1108	TACATTGTGCATGGAGCACGTGACGAGGCAACCTGGTGAACCTTCTGCGGACCCGGGGT	1166
QY	321	ArgAlaIleuValAsnThrAlaGlnIleuIleuGlnPheSerIleuHisValAlaGlnGlyMet	340

```

Db      1168 CGAGCCCTGCTGACACCCGCTAGCTCTGACAGTTTCTCTGACAGCTGCCGAGGGCAGT 1227
QY      341 GUTYrLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
      |||||
Db      1228 GAGTACTGTGAGGAGCAAGACTTGTGACACCGGAGCTGGCCCGGCAACATCTCGTGC 1287
QY      361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
      |||||
Db      1288 TCAGAGAGCTGTGGTGGCCAAAGGTACAGCACTTGGCCCTGGCCAAAGCCGAGGGAAGGGG 1347
QY      381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
      |||||
Db      1348 CTAGACTCAAGCCGGGTGCTGCCCTCAAGTGGAGCGCCGCCGAGGCTCTCAAAACAGGGGAAG 1407
QY      401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyr 420
      |||||
Db      1408 TTCACGACGACATCGGATGTGTGAGACTTTGGGGGTCTGCTCTGGGAGGCTCTTCATAT 1467
QY      421 GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
      |||||
Db      1468 GGACGGGCTCCGTACCCCTAAATGTCATGAAAGAGGTGTGGAGGCCGTGGAGAGAGGGG 1527
QY      441 TYrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerSerCys 460
      |||||
Db      1528 TACCGATGGAACCCCGCCAGGGGCTGTCCAGGGCCCGCTGACGCTCTCATGAGCACTGC 1587
QY      461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
      |||||
Db      1588 TGGGAGGCGAGAGCCCGCCCGCCGACCTTCCGAAACCTGGCGAGAGAGCTGGCCGG 1647
QY      481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGluAspAlaAspGlySerThr 500
      |||||
Db      1648 GAGCTACGACAGTCAGAGTGTCCCAAGCTCTGCTCAGGCGAGAGACCCGACGCTCCACC 1707
QY      501 SerProArgSerGluPro 507
      |||||
Db      1708 TCGCCCCGAGACGACGAGGCC 1728

```

```

DR      P-PSDB; AAR84181.
XX
XX      New poly:nucleotide(s) encoding megakaryocyte tyrosine kinase(s) -
PT      used to develop prods. for the treatment and diagnosis of kinase
PT      related signal transduction abnormalities.
XX
XX      Claim 2: Fig 1A-C; 82pp; English.
XX
XX      Overlapping cDNA clones were combined to obtain a sequence (AAT00616)
CC      coding for human megakaryocyte kinase MKK1 (AAR84181). The clones
CC      were isolated from a fetal brain library using degenerate primers
CC      (AAT00614-15) based on conserved regions within the kinase domain of
CC      receptor tyrosine kinases. MKK polynucleotides can be used in the
CC      prodn. in host cells of recombinant MKK, and in the gene therapy of
CC      diseases such as acute megakaryocytic leukaemia, myelofibrosis and
XX      acute megakaryocytic myelosis.
SQ      Sequence 2000 BP; 375 A; 631 C; 674 G; 320 T; 0 other;
Alignment Scores:
Pred. No.: 2,1e-198 Length: 2000
Score: 2668.00 Matches: 506
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.80% Mismatches: 0
Query Match: 99.89% Indels: 0
DB: 16 Gaps: 0
US-09-977-260-2 (1-507) x AAT00616 (1-2000)
QY      1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
      |||||
Db      258 ATGGGGGGGAGAGGCTCTGTGGTTCCTGGCGGCAATTCACAGGCTGTGATCTGTGAG 317
QY      21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
      |||||
Db      318 GAACCTTCCCGGGGTAGAGCCCGCTTCTCCAGAGCTGGACACCCCTCCGCTCAGGC 377
QY      41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
      |||||
Db      378 AGGATGCCAAGCAGAGCGGTGGGCCCGCGGACCCAGCTGATACCAAAATGCAGACACACC 437
QY      61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValValThrIleLeuGluAla 80
      |||||
Db      438 CGCCCCCAAGCCAGGAGAGCTGTCCGCAAGGGGAGAGCTGTGCACATCTCGAGAGGCC 497
QY      81 CysGluAsnLysSerTrpTrpArgValLysHisIleThrSerGlyGlnGluLeuLeu 100
      |||||
Db      498 TCGGAGAACAAAGCTGTACCGCGTCAAGACACACACAGTGGAGAGAGGCGCTGCG 557
QY      101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
      |||||
Db      558 GCAGCTGGGGGCGCTCGGGAGCGGAGGCCCTCTCCGACCCCAAGCTCAGCTCATG 617
QY      121 ProThrPheHisGlyLysIleSerGlyGlnGluAlaValGlnGluLeuLeuProProGlu 140
      |||||
Db      618 CCGTGGTTCACAGGGAAGATCTCGGGCCAGAGGCTGTCTCAACACCTTCAAGCTCCCGAG 677
QY      141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysVal 160
      |||||
Db      678 GATGGGCTGTCTCTGTGGGAGATCCGGGCGCCACCCCGGAGACTCTCTGTGCGG 737
QY      161 SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisIleuThrIle 180
      |||||
Db      738 AGCTTTGGCGCGAGCTGATCCATACCGCGCTGTGCACCCGAGAGGCGACCTCAAAATC 797
QY      181 AspGluAlaValPhePheCysAsnLeuMetIlePheMetValGlnHisTyrSerLysArgLys 200
      |||||
Db      798 GATGAGGCGCGTGTCTGTGCAACTCATGACATGACATGTGACATTAACAGCAAGGACAAAG 857
QY      201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
      |||||
Db      858 GCGCGTATCTGACCAAGCTGTGTGAGACCAAGCGGAAACAGCGGACCAAGTGGCGGAG 917

```


QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValValThrIleLeuGluAla 80
 Db 443 CGCCCCAAGCCAGGGAGCTGGCTTCCGCAAGGGCCAGCGTGCACATCTCGAGGCC 502
 QY 81 CysGluAsnLysSerTrpTyrArgValLysHisHisThrSerGlyGlnGluGlyLeu 100
 Db 503 TGCAGAACACAGACTGCTGACCGCGTACAGACACACAGTGGACAGAGGGCTCTCG 562
 QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 Db 563 GCACCTGGGGCGCTGGCGGAGCGGCCCTCTCCGACAGCCCAAGCTCAGCTTATG 622
 QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
 Db 623 CCGGGTTCACAGGAGATCTCGGCGAGAGCTGTCCAGCAGCTGACGCTCCGAG 682
 QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuGlyVal 160
 Db 683 GATGGCTGTCTCGTGGCGAGTCCGCGCCACCCCGCGCACTGCTGTGGCTG 742
 QY 161 SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle 180
 Db 743 AGCTTGGCCGCGACGTCATCCATCCGCTGCTGACCGCGCCACCTCAGCATC 802
 QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLysAspLys 200
 Db 803 GATGAGCCCGTGTCTTCTCAACCTCATGAGATGTTGAGATACACAGAGACAG 862
 QY 201 GAlaValIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerLysAlaGlu 220
 Db 863 GCGCCTATCTGCACCAAGCTGTGTAGACCAAGGGGAAACCGGAGCAAGTGGCCGAG 922
 QY 221 GluGluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
 Db 923 GAGAGCTGGCCAGGCGCGCTGTACTGACCTGACATTTGACATTTGGAGCCAG 982
 QY 241 IleGlyGluGlyGluPheGlyAlaValLeuGlnGlyGlyTyrLeuGlyGlnLysValAla 260
 Db 983 ATCGAGAGGGAGAGTGTGAGCTGTCTCAGAGGTGATGATCTGGGGCAAAAGTGGCC 1042
 QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGlyThrAlaValMet 280
 Db 1043 GTGAGAAATTCAGATGTGATGTGACAGCCAGGCTTCTGTGAGAGAGCGCGTATG 1102
 QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuGluGlyAlaIleLeuHisGlnGlyLeu 300
 Db 1103 ACGAAGATGCAACACGAGAACGTCGTGCTGTGGCGCTGATTCGACACAGGGGCTG 1162
 QY 301 TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
 Db 1163 TACATTGTCTATGAGACGTCGACCAAGGGCAACCTGTGTAACCTTTCGCGAGCCGGGCT 1222
 QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlnGlyMet 340
 Db 1223 CGAGCCCTGTGAAACACCGCTCACTGCTGTGACAGTTCCTGTGACGCGGCGAGGCGATG 1282
 QY 341 GluTyrLeuGluSerLysLysLeuValHisArgAspLeuAlaIlaArgAsnIleLeuVal 360
 Db 1283 GAGTACTGTGAGAGCAAGAACTTGTGACCGGAGCTGGCCCGCGCAACATCTGTGTC 1342
 QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGlnArgLysGly 380
 Db 1343 TCAGAGAGACTGTGTGCGCAAGGTCACGACTTGTGGCTGTGCGCAAGCGCGAGAGGGG 1402
 QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
 Db 1403 CTGAGACTCAAGCCGCGCTCCGCTCAAGTGACGCGCGCGAGGCTCTCAAAACCGGG--- 1459
 QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyr 420
 Db 1460 TTACACAGCAAGTCGATGTCTGAGATTTTGGGGTCTGTCTCTGAGGAGGTCTTCTCATAT 1519

QY 421 GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
 Db 1520 GAGCGGCTCCGTCACCTAAATGTCACTAAAGAGGTGTGCGAGCGCTGAGAAAGGGG 1579
 QY 441 TyrArgMetGluProProGluGlyCysPheProGlyProValHisValLeuMetSerSerCys 460
 Db 1580 TACCGCATGGAAACCCCGGAGGGCTGTCAAGGCCCCCTGTGACGCTCATGACACAGCTGC 1639
 QY 461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
 Db 1640 TGGGAGGCACAGCCGCC-CCCGGCCACCTTCGCGAAACTGCGCGAGAAAGCTGGCCCG 1698
 QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
 Db 1699 GAGCTACGACAGTGTGAGGTGCCCGACCTCGCTCAAGGCGACGAGCGCGAGGTC-ACC 1757
 QY 501 SerProArgSerGlnGluPro 507
 Db 1758 TCGCCCCGAAAGCAGAGGCC 1778

RESULT 5
 ABR88791
 ID ABR88791 standard; cDNA; 1713 bp.
 AC ABR88791;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE cDNA encoding human kinase related to tyrosine kinase family.
 XX
 KW Human; kinase; tyrosine kinase; kinase mediated disorder;
 KW hematopoietic disorder; cell signal transduction disorder;
 KW cancer; haemostatic; cytosolic; chromosome 19; gene; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 223..1470
 FT /*tag= a
 FT /*product= "kinase"
 XX
 PN W0200252018-A2.
 XX
 PD 04-JUL-2002.
 XX
 PF 19-DEC-2001; 2001MO-US48546.
 XX
 PR 21-DEC-2000; 2000US-0741154.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Beasley EM, Shao W, Ketchum K, Di Francesco V;
 XX
 DR WPI: 2002-583568/62.
 DR P-PSDB; AAU99575.
 XX
 PT New isolated human kinase proteins and genes, useful in developing
 PT drugs, as well as for diagnosing, preventing or treating disorders
 PT associated with defective cell signal transduction, e.g. cancer or
 PT haematopoietic disorders
 XX
 PS Claim 4; Fig 1A; 78bp; English.
 XX
 CC The present invention relates to the isolation of a novel human kinase
 CC related to the tyrosine kinase family, and polynucleotide sequences
 CC encoding it. The gene encoding the kinase of the invention maps to
 CC chromosome 19. The kinase may be used for identifying a modulator of
 CC the kinase, an agent that binds to the kinase, or for identifying
 CC other members of the family. The kinase may also be used to raise
 CC antibodies which may be used in immunoassays or drug screening assays.
 CC The sequences of the invention may be used for treating a disease or
 CC condition mediated by a human kinase such as haematopoietic disorders
 CC and other disorders associated with defective cell signal transduction

CC (e.g. cancer). The polynucleotide sequences encoding the kinase are
 CC useful for isolating and purifying the kinase, and as probes or primers.
 CC The present sequence encodes the human kinase of the invention.

XX
 XX Sequence 1713 BP; 374 A; 499 C; 575 G; 265 T; 0 other;

Alignment Scores:

Score:	4.94e-148	Length:	1713
Percent Similarity:	2020.50	Matches:	410
Best Local Similarity:	84.88%	Conservative:	11
Query Match:	82.66%	Mismatches:	48
	75.65%	Indels:	29
		Gaps:	6

US-09-977-260-2 (1-507) x ABK88791 (1-1713)

```

OY 23 ProArgValSerPro-----ArgPheLeuAlaGlyAlaTrp 33
    |||||
DB 30 CCCCAGAGAGAGCGGGGGTGGCAGAGAACACACAGGCTGCAGGGGGCTGGAGAACAA 89
    |||||
OY 34 HisProProProValSerAlaArgMetProThrArgArgTrpAlaProGlyThrGlnCys 53
    |||||
DB 90 CACCCCTCTCTCC-----TCCCCCAGAGAACTGGACCGGGGACCTCG--- 134
    |||||
OY 54 IleThrLysCysGlnHisThrArgProLysProGlyIleuAlaPheArgLysGlyAsp 73
    :|||
DB 135 GTTGGCGA-TGC-----TGGCGGCGAGCCGG 160
    |||||
OY 74 ValValThrIleLeuGlnAlaCysGlnLysSerTrpTrpArgValLysHisHisThr 93
    :|||
DB 161 TTGCTAGCGCATGCGAGAGAGAGAGAGAAAGTCTTCATGGGGTCTAGAGCATCTT 220
    |||||
OY 94 SerGlyGlnGlnGlyLeuLeuAlaAlaGlyAlaLeuArgLysGlnAla---LeuSer 112
    |||||
DB 221 AAATG-TCCCTCTCATGAGGACCATTTTCCCTGTAAGAGAGAGAGAGACCTCCGCC 279
    |||||
OY 113 AlaAspProLysLeu---SerLeuMetProTrpPheHisGlyLysIleSerGlyGln 131
    |||||
DB 280 GAGGAGACCCGTCGTGACGACGACGCTTGTGGTTCACAGGGAGATCTCGGGCCAGAG 339
    |||||
OY 132 AlaValGlnGlnLeuGlnProProGlnAspGlyLeuPheLeuValArgGlnSerAlaArg 151
    |||||
DB 340 GCTGTCCAGACACTGACGCTCCAGGATGGCTGCTCTGTCGCGGAGTCCCGGCC 399
    |||||
OY 152 HisProGlyAspTrpValLeuLysValSerPheGlyArgAspValIleHisTrpArgVal 171
    |||||
DB 400 CACCCCGGCTAGCTCTGTCGTGAGCTTTGGCCGAGCTATCCACTACCGCGTG 459
    |||||
OY 172 LeuHisArgAspGlyHisLeuThrIleAspGlnAlaValPhePheCysAsnLeuMetAsp 191
    |||||
DB 460 CTGCACCGGAGCGGCACTCAACATCGATGAGCGCGTCTTCTTGCACCTCATGGAG 519
    |||||
OY 192 MetValGlnHisTrpSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLys 211
    |||||
DB 520 ATGGTGGACATATACAGCAAGGACGAGGCGCTATCTGCACCAAGCTGTGAGACCAAG 579
    |||||
OY 212 ArgLysHisGlyThrLysSerAlaGlnGlnGlnLeuAlaArgAlaGlyTrpLeuLeuAsn 231
    |||||
DB 580 CGGAAACACGGGACCAAGTGGCGGAGAGAGACTGGCCAGCGGCGGTGTTACTGAAAC 639
    |||||
OY 232 LeuGlnHisLeuThrLeuGlnGlyAlaGlnIleGlyGlnGlyLeuPheGlyAlaValLeuGln 251
    |||||
DB 640 CTGCACATTTGACATTTGGGAGCACAAGATGGAGAGGAGATTGGAGCTGTCTGCAG 699
    |||||
OY 252 GlyGlnTrpLeuGlnGlnLysValAlaValLysAsnIleLysCysAspValTrpAlaGln 271
    |||||
DB 700 GGTGAGTAACTGGGGGCAAAAGTGGCGGTGAAGAAATATCAAGTGTATGTGACAGCCAG 759
    |||||
OY 272 AlaPheLeuAspGlnTrpAlaValMetThrLysMetGlnHisGlnLysLeuValArgLeu 291
    |||||
DB 760 GCGTCTCTGAGCAGAGCGCGCTCATATACCAAGATGCAACACAGAACCTGGTGGCTCTC 819
    |||||
OY 292 LeuGlyValIleLeuHisGlnGlyLeuTrpIleValMetGlnHisValSerLysGlyAsn 311
    |||||

```

```

DB 820 CTGGCGGTGATCTGCACACAGGGGGTGTACATTTGATGAGAGACCTGGAGGCAAC 879
    |||||
OY 312 LeuValAsnPheLeuArgTrpArgGlyArgAlaLeuValAsnThrAlaGlnLeuGln 331
    |||||
DB 880 CTGTGAACTTTTGGGGACCGGGTCCAGCCCTGTGTAACACCCCTCAGCTCCCGAG 939
    |||||
OY 332 PheSerLeuHisValAlaGlnGlyMetGlyTrpLeuGlnSerLysLeuValHisArg 351
    |||||
DB 940 TTTTCTCTGACCTGGCGGAGGCGCATGAGATACCTGGAGAGCAAGAAAGCTTGTGACACCG 999
    |||||
OY 352 AspLeuAlaAlaArgAsnIleLeuValSerGlnAspLeuValAlaLysValSerAspPhe 371
    |||||
DB 1000 GACCTGGCCCGCCGCAACATCTGTCTCAGAGAGCTGGTGGCCAAAGTCAAGGCTTT 1059
    |||||
OY 372 GlyLeuAlaLysAlaGlnArgLysGlyLeuAspSerSerTrpLeuProValLysTrpThr 391
    |||||
DB 1060 GGCCTGGCCAAAGCCCGAGGAGGAGGCTGAGCTCAAGCCGGCTGCCGTCAGTGGACG 1119
    |||||
OY 392 AlaProGlnAlaLeuLysHisGlyLysPheThrSerLysSerAspValTrpSerPheGly 411
    |||||
DB 1120 GCGCCGAGGCTTCAAACACAGGAGAGTTCACCGCAAGTCGATGTCGTGAGATTGGG 1179
    |||||
OY 412 ValLeuLeuTrpGlnValPheSerTrpGlyArgAlaProTrpProLysMetSerLys 431
    |||||
DB 1180 GTCTGCTCTGGAGGCTTCTCATATGACGGGCTCGTACCTTAAATGTCACGTGAA 1239
    |||||
OY 432 GlyValSerGlnAlaValGlnLysGlyTrpArgMetGlnProProGlnGlyCysProGly 451
    |||||
DB 1240 GAGGTGTCGAGGCGCTGAGAAAGGAGTACCGCATGGAACCCCGAGGCTGTCAGAGC 1299
    |||||
OY 452 ProValHisValLeuMetSerSerCysTrpGlnAlaGlnProAlaArgArgProPhe 471
    |||||
DB 1300 CCGGTCGACGTCTCATGACGACGCTGAGGAGGACAGGCGCCCGCCGACCCCTTC 1359
    |||||
OY 472 ArgLysLeuAlaGlnLysLeuAlaArgGlnLeuArgSerAlaGlyAlaProAlaSerVal 491
    |||||
DB 1360 CGCAACATGGCGGAGAACTGGCCGGAGACTACGACATGCAAGTGGCCCGACCGCTC 1419
    |||||
OY 492 SerGlyLysAspAlaAspLysSerThrSerProArgSerGlnGlnPro 507
    |||||
DB 1420 TCAGGCGAGAGAGCGGCGCTCCACCTCGCCCGCAAGACGAGACCC 1467
    |||||

```

RESULT 6
 AB199884
 ID AB199884 standard; cDNA; 8212 BP.
 AC AB199884;
 XX
 XX 07-MAR-2002 (first entry)
 DT
 XX
 XX Mouse ischemic condition related cDNA sequence SEQ ID NO:1027.
 DE
 XX
 XX Mouse; ischemia; compressive ischemia; occlusive ischemia;
 KW vasospastic ischemia; ischemic condition; ischemic disease; ss.
 XX
 XX Mus musculus.
 OS
 XX
 XX MO200188188-A2.
 PN
 XX
 XX 22-NOV-2001.
 PD
 XX
 XX 18-MAY-2001; 2001MO-IP04192.
 PF
 XX
 XX 18-MAY-2000; 2000JP-0145977.
 PR
 XX
 XX (UYNT-) UNIV NITHON SCHOOL JURIDICAL PERSON.
 PA
 XX
 XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
 PI
 XX
 XX WPI: 2002-034733/04.
 DR
 XX
 XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring

PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
genes -

XX PS Claim 2: Page 2592-2596; 2690pp; English.

CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (1) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (1). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (AB199202 to AB199912, encoding
CC the protein sequences in AB199202 to AB199912) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. AB199913 and AB199914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.

XX Sequence 8212 BP; 1663 A; 2301 C; 2300 G; 1948 T; 0 other:

Alignment Scores:

Pred. No.: 9, 21e-110 Length: 8212
Score: 1540.50 Matches: 420
Percent Similarity: 34.26% Conservative: 23
Best Local Similarity: 32.48% Mismatches: 39
Query Match: 57.68% Indels: 816
DB: 24 Gaps: 12

US-09-977-260-2 (1-507) x AB199884 (1-8212)

QY 26 SerProArpPheLeuArGAlaTrpHisProProValSer-----AlaArg 41
Db 4184 ACCCCAGCTTCTCTCAGAGTCTCTGTCCTCCCAATTCACACTTGCCTGTCTCA 4243
QY 42 MetPro-----ThArGArGTTPAlaProGlyThrGln 52
Db 4244 GTCCTGACATACCTCTGCTCTGTCCTCCAGACGCTGGGCCCCCTGGACTCAA 4303
QY 53 CysIleThrLysCysGlnHisThrArGProLysProGlyGluLeuAlaPheArgLysGly 72
Db 4304 TCATGACCAAGTGTGAGACTCTGCCCCAAGCCGGGTGAGCTACCTTTGAAAGGCT 4363
QY 73 AspValValThrIleLeuGluAlaCysGluAsn----- 83
Db 4364 GACATGCTGACCATCTTGGAGGCTGTAGGT-GAGAGGGGAGCTGGAGGTAGGTGTG 4422
QY 83 ----- 83
Db 4423 TGCCTTGGGGTGCCAGGGCCGACACCACAGCTCTTCCCACATCTGTCCCTTCT 4482
QY 84 -----LysSerTriPtyrArgValLysHisHisThrSerGlyGlnGluLeu 100
Db 4483 GTACAGACACAGACTGTGACGAGCCAGCACACATGCGAGTGGGAGGAGGCTGTG 4542
QY 101 AlaAlaGlyAlaLeuArGlnuArgLysAlaLeuSerAlaAspProLysLeuSerLeuMet 120
Db 4543 GGGCGCGCTCTGTGAGACAGCGGGAGCCCTCTCCACAGACCCCAAGCTCAGCTCAG 4602
QY 121 Pro----- 121
Db 4603 CCGTAGTGGCACCTGGGGGGTGTGTGAGAGGCTCAGCCCCCTTACCCCTGCTTAC 4662
QY 122 -----TrpPheHisGlyLysLysLysSerGlyGlnGlu 131
Db 4663 AAGAGCCCTCTTCACACACACATCCCTACGATGTGTTTCATGGAAGATTCGCGCCAGA 4722
QY 131 uAlaValGlnGlnLeuGlnProProGluuAspLysLeuPheLeuValArgGluSerAlaAr 151
Db 4723 AGCCATACAGAGCTGACAGCCACCGAGAGAGCGGCTGTCTTGTGTGAGGGAATCAGCCG 4782

QY 151 gHisProGlyAspTyrValLeuCysValSerPheGlyArGAspValIleHisTyrArgVa 171
Db 4783 TCACCTGGAGACTATGCTGTGTGTGCTGAGTTTGGCCGTGACGTCATCCACTACGCTGT 4842
QY 171 IleuHisArGAspGlyHisLeuThrIleAspGluAlaValIlePhePheCysAsnLeuMet 191
Db 4843 TTGCAATCAGATGGGACACTCACCATCATGATGAGGCGGTGTCTGTAACTGATGA 4902
QY 191 pheValGlu----- 194
Db 4903 CATGTGTGA-GGTGACGTGTGGGCAAGGCTGGGAGAGGAGGATGAAATATGATGTGT 4961
QY 194 ----- 194
Db 4962 GGCCCGAGGAACTGTCACTTCTTCTGTAGAACTGGGGTGCACCTCCCGAGGAGCTTAT 5021
QY 194 ----- 194
Db 5022 GTTCACACCAACACACCCCTCAAGTTCCATTTCAACCCCACTGTTGCCATGAACCC 5081
QY 194 ----- 194
Db 5082 GATCACCTTAGCTGTGCACTTAGAGAGAGAGTGTGCAACTCTCACTGAACCTTGGGGTGG 5141
QY 195 -----HisTyrSer 197
Db 5142 GATGTGGGAGCGGTCTCTGACAGGAGAGCTCATATGGCGGGCTCCACACACTACAC 5201
QY 197 rLysAspLysGlyAlaIleCysThrLysLeuValArGProLysAsArgLysHisGlyThrLys 217
Db 5202 CAAGACCAAGGGGGCCATCTGCAACCAAGCTGTGAAGCCAGAGGAGAAACAGCGCAAA 5261
QY 217 sSerAlaGlnGlnGluLeuAlaArg----- 225
Db 5262 GTCTGAGAGAGAGAGAGCTGCCAA-GGGATAGAGAGCCATAGTGAGAGAGAGATCT 5320
QY 225 ----- 225
Db 5321 CGTGGCAGACAGAGAGACTCCACCTCCAGACTCAGACCCCTAATTCATTGCTACCC 5380
QY 225 ----- 225
Db 5381 TCACACCCAGCTTGGCTTGTGTTAAGCCCTTGCCTTCCAACTGTGACACTTATCC 5440
QY 225 ----- 225
Db 5441 ACTGGGGGCTATGATCATCAGAGCAAAATTCGAGTGGAGAGGTGTCTGTCTGCTTGG 5500
QY 225 ----- 225
Db 5501 AATCTTACAGAGTGGCGGTAGGGTGGGGGTTAGAGGTCTTTGGAAGGAGAGGAGGCTTTC 5560
QY 225 ----- 225
Db 5561 AGAGCAACTGCAGACTCTGTGAGAGTACAGATTAAGAGACTATAGCCCTGAAGAATY 5620
QY 225 ----- 225
Db 5621 CCAGAACCTGTGCAAACTCTGTATTTCTGTGCTTGTGTCTGTGTGGGTGAGCAAG 5680
QY 225 ----- 225
Db 5681 GGGAGGCCCTTCCAGAGGCAATGGGGGCGACAGGGGAGTCACTCTGTCCTTAGA 5740
QY 225 ----- 225
Db 5741 GGGAGGCTCTGAGCATGAGCTCAAGTATAGGAACTTAGGGGGGAGCTGTATTCTGATGA 5800
QY 225 ----- 225
Db 5801 AGCGTGGATGTGCTGTGCTTGCAGAGGCTGTAAATATAGTACAGAGAGTGGGTGTCTTCA 5860
QY 225 ----- 225

Db 5661 TGCAGCCTTAGGCGAGCTGTGTGAGACAGGACATGCTCAGGTGTGCCAGGCTGTGGGAT 5920
Qy 225 -----
Db 5521 ATCAGCTCTCCGAGCTCTCAGTGAATGCMAACTGGGTGGGATGGCGAGGAGTGCTCTGCC 225
Qy 226 -----
Db 5981 CTCCTCTCAGTGGGCTCTTGTCTGTCTGACAGCTGGGCTACTGCACTGCACTGCACTCT 5980
Qy 235 -----
Db 5981 CTCCTCTCAGTGGGCTCTTGTCTGTCTGACAGCTGGGCTACTGCACTGCACTGCACTCT 6040
Qy 235 uThLeuclYalaglInlelYgluglYgluPhelYala----- 248
Db 6041 GACTCTGGAGCACAGATTGGAGAGGGAGGAGTTGGAGG-TGAGAGGGGACACAGCAGG 6099
Qy 248 -----
Db 6100 CAGTGTTCAGAAATTGACCTTCAACTGACAGTGACAGAGACAGGTGGCTGTGGGA 6159
Qy 248 -----
Db 6160 GTACGCACACCCCTCAGCTCCAGGATCCTTATTCAGAAACCTTCCCTGTGTGTGT 6219
Qy 248 -----
Db 6220 TCTGGGCTTGGGTACAGAGACAGTAGACTTCAGGATGTCTCAGTGCAGCCGCA 6279
Qy 248 -----
Db 6280 AATTATGCMAACATGGCGACGACACCTGTGGTCTTCATGAGACAGAAATTTGATTCGT 6339
Qy 248 -----
Db 6340 AATGAGACAGCATGTATTTCAGATGATTAATTTTTCCTCATTCAGGGCTGGGG 6399
Qy 248 -----
Db 6400 GATGAGACAGGCTCTGTAATGATTTTCAACTCTGTGCCAGGCGACACCCCGAGCC 6459
Qy 248 -----
Db 6460 CTCAGTGGGGATTCCAGGAGGGGTGTGACAGAGGTGTGTCTGTCTGCTGCTGCTG 6519
Qy 249 -----
Db 6520 TGGGCGCTGACCTGTCTCTCTCCACAGCGCTCTACAGGGGTGAGTCTGGGACAGAA 6579
Qy 258 sValAlaValLysAsnLeuLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAl 278
Db 6580 GGTGGCTGTGAAGATATCATGATGTGACAGCCAGGCTTCTGATGAGACGCG 6639
Qy 278 aValMetThr----- 281
Db 6640 TGTATGACGTGAGTGTGGGGTGGGCCAGGCTCCAGGCTGAGGTCCAGGTTATAC 6699
Qy 282 -----
Db 6700 TCACATCATGTCCCTCTGATCCCTGACACCCAGAGACCTGAGAGACAGAACTGAG 6759
Qy 289 aLaArgLeuLeuGlyValIleLeuHisGlnGlyLeuTyrIleValMetGlnHisValSerL 309
Db 6760 TGCACATCTCTGGGTGTGATCTGCACACGCGCTGTACATGTGATGAGACAGTGAACA 6819
Qy 309 ys----- 309
Db 6820 AGGTGTGCAGGGGTGGGTGGGTGGGACAGAAAGGTGGCCCAACAGCTGAGACACCC 6879
Qy 310 -----
Db 6880 CAAGTCCCTCCCTTAACAGAGGCACTGTGTAACCTCTGCGAGACGCGGGCGCTGCT 6939
Qy 323 LeuValAsnThrAlaGlnLeuLeuGlnPhe----- 332

Db 6940 CTGTGAGCACCTCTCAGCTTCTGTCAGTTGTGCTGTGAATGATTCATCCATCCAGAGGT 6999
Qy 333 -----
Db 7000 AGTCCCAAGGTGAAGGAAACCTGAGGCTCTGACTTCAAGGTGAGCCGCTTGTGTCT 7059
Qy 334 Leu--HisValAlaGlnGlyMetGluTyrLeuGlnSerLysLysLeuValHisArgAspL 353
Db 7060 CTAGTCACTGTGCTCAAGAGCATGGAATACCTGAGAGAGCAAGAGCTGTGACACCGGACC 7119
Qy 353 euaLaAlaArgAsnLeuValSerGlnAspLeuValAlaLysValSerAspPheGlyL 373
Db 7120 TGGCTGTGGGAACATCTGCTGTGTGAGACTTGTGGCCAAAGGTCAGTCACTTGGCT 7179
Qy 373 euaLaAlaGlnAlaGlnArgLysGlyLeuAspSerArgLeuProValLysThrPheAlaP 393
Db 7180 TAGCCAGGCGAGAGCGCAAGGGGCTGAGCTCAAGCGGCTCCAGTCAAGTGAAGGAGC 7239
Qy 393 roGluAlaLeuLysHisGly----- 399
Db 7240 CTGAGCTCTCAAAAACGGGGGTGAGCAGCCTTCTACCTCTCTGGGATTAAGGGTTCC 7299
Qy 399 -----
Db 7300 TCAGAGGCGGCTGACGTTCTTGATTTGCTCTGACATATGACACCCCTCAACCTC 7359
Qy 399 -----
Db 7360 TCTGAGCTCAAGGGGAAATGTTGCTCTCCAGTGAAGCCCTAAGCTCAACAATTCCTC 7419
Qy 399 -----
Db 7420 CTAGACTTCCCTCAGCTGATTCCTCCATCTCTGAGAGAGGCTGGCTTGGACAGAGG 7479
Qy 399 -----
Db 7480 AGGAGATGAGCTGGGATTTTCCAGGAGGCGCAGAGGGGCTCTCACCCTGACCTC 7539
Qy 400 -----
Db 7540 TGCCTCAACAGCGGTTCTCCAGCAGTGGATGTGAGTTGGGGTGTGCTGTGGG 7599
Qy 416 LuValPheSerTyrIleValAlaProTyrProLysMetSer----- 429
Db 7600 AAGTCTTCTTATGAGAGAGCCCATACCCCAAGATGTG-TGACAGCACCTTGAAGGAG 7658
Qy 429 -----
Db 7659 GAGGCGAAGACAGACATTCAAGCAGCTGTGTGACATGTCTGCCACATGAACCTCT 7718
Qy 429 -----
Db 7719 GAGCAGCCCTGATGTTGATGTCTGTATGACAGCTGTGTATCTGTCTCACTGACTGCC 7778
Qy 430 -----
Db 7779 CTCCAACCCAGGCTGACACCCCTGTGTCCACAGTGCCTAAAGAGGTTTCAAGAGCTGTGG 7838
Qy 438 LuLysGlyTyrArgMetLupProPogluGlyCysPProGlyProValHisValLeuMetS 458
Db 7839 AGAAGGCTTACCGCATGAGAGCCCGGATGGCTGCCAGGCTCTGTGCACACCCCTCAAGG 7898
Qy 458 erSerCysTrpGlnAlaGluProAlaArgArgProPheArgLysLeuAlaGlnLysL 478
Db 7899 GTAGCTGTGGAGAGGCAAGGCTGGGCGCCAGCACCTTCCGCAAAATAGTGGGAAGC 7958
Qy 478 euaLaArgLupLeuArgSerAlaGlyAlaProAlaSerValSerLysGlnAspAlaAsp 498
Db 7959 TGGGCGTGTGAGCTGTGAGTGTCTCGGCGCCGCTGAGGAGACAGAGGCTGTAGG 8018
Qy 498 LysThrThrSerProArgSerGlnLupPro 507
Db 8019 GCTCAGCTCTCCACAGGAGGCGAGCCAGCC 8047

Feature	Location/Qualifiers
Key	3000..14181
CDS	/tag= a /product= "Kinase" 3000..3085
exon	/tag= b /number= 1 3086..8330
Intron	/tag= c /number= 1 8331..8550
exon	/tag= d /number= 2 8551..9144
Intron	/tag= e /number= 2 9145..9238
exon	/tag= f /number= 3 9239..10692
Intron	/tag= g /number= 3 10693..10758
exon	/tag= h /number= 4 10759..12566
Intron	/tag= i /number= 4 12567..12666
exon	/tag= j /number= 5 12667..12746
Intron	/tag= k /number= 5 12747..12831
exon	/tag= l /number= 6 12832..12912
Intron	/tag= m /number= 6 12913..12986
exon	/tag= n /number= 7 12987..13176
Intron	/tag= o /number= 7 13177..13372
exon	/tag= p /number= 8 13373..13768
Intron	/tag= q /number= 8 13769..13855
exon	/tag= r /number= 9

FT	intron	13856..13941	
FT		/*tag- s	
FT		/number= 9	
FT	exon	13942..14178	
FT		/*tag- t	
FT		/number= 10	
FT	variation	(replace 1130, A)	
FT		/*tag- u	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 4245, A)	
FT		/*tag- v	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 4248, T)	
FT		/*tag- w	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 4250, A)	
FT		/*tag- x	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 4255, A)	
FT		/*tag- y	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 4258, T)	
FT		/*tag- z	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 4260..4261, T)	
FT		/*tag- aa	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 4266..4267, T)	
FT		/*tag- ab	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 4269..4270, T)	
FT		/*tag- ac	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 4273, T)	
FT		/*tag- ad	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 4275, A)	
FT		/*tag- ae	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 4401, T)	
FT		/*tag- af	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 4551..4552, C)	
FT		/*tag- ag	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 4739, G)	
FT		/*tag- ah	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 6309, A)	
FT		/*tag- ai	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 8753..8754, T)	
FT		/*tag- aj	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 10797, T)	
FT		/*tag- ak	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 10806, G)	
FT		/*tag- al	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 11199, T)	
FT		/*tag- am	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 11201, T)	
FT		/*tag- an	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 11319, C)	
FT		/*tag- ao	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 12238, C)	
FT		/*tag- ap	
FT	variation	/standard_name= "Single nucleotide polymorphism (SNP) "	
FT		(replace 14139, T)	

```

FT FT /*tag= aq
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT (replace 15375, C)
FT FT /*tag= ar
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX PN MO200252018-A2.
XX PD 04-JUL-2002.
XX PF 19-DEC-2001; 2001WO-US48546.
XX PR 21-DEC-2000; 2000US-0741154.
XX XX (PEKE ) PE CORP NY.
XX PI Beasley EM, Shao W, Ketchum K, Di Francesco V;
XX DR WPI: 2002-583568/62.
XX DR P-PSDB; AAU99575.
XX XX
XX PT New isolated human kinase proteins and genes, useful in developing
XX PT drugs, as well as for diagnosing, preventing or treating disorders
XX PT associated with defective cell signal transduction, e.g. cancer or
XX PT hematopoietic disorders
XX PS Claim 4; Fig 3A-F; 78pp; English.
XX XX
XX CC The present invention relates to the isolation of a novel human kinase
XX CC related to the tyrosine kinase family, and polynucleotide sequences
XX CC encoding it. The gene encoding the kinase of the invention maps to
XX CC chromosome 19. The kinase may be used for identifying a modulator of
XX CC the kinase, an agent that binds to the kinase, or for identifying
XX CC other members of the family. The kinase may also be used to raise
XX CC antibodies which may be used in immunoassays or drug screening assays.
XX CC The sequences of the invention may be used for treating a disease or
XX CC condition mediated by a human kinase such as hematopoietic disorders
XX CC and other disorders associated with defective cell signal transduction
XX CC (e.g. cancer). The polynucleotide sequences encoding the kinase are
XX CC useful for isolating and purifying the kinase, and as probes or primers.
XX CC The present sequence encodes the human kinase of the invention.
XX XX
XX SQ Sequence 16389 BP; 3511 A; 4901 C; 4515 G; 3462 T; 0 other;
XX
Alignment Scores:
Pred. No.: 7.56e-104 Length: 16389
Score: 1469.50 Matches: 501
Percent Similarity: 21.59% Conservative: 0
Best Local Similarity: 21.59% Mismatches: 6
Query Match: 55.02% Indels: 1819
DB: 24 Gaps: 12
US-09-977-260-2 (1-507) x ABK88792 (1-16389)
OY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
DB 7229 ATGCGGGGGGAGGCTCTGTGTCTGCGGGGCGCATTCAGCGCTGATTCGTGAG 7288
OY 21 Glu----- 21
DB 7289 GAATCTCCCGGGTAAGATCACTTCCCAACAGGGGCTTGGGAGCCAGTTCTGGGTCCC 7348
OY 21 ----- 21
DB 7349 ATCCAGATGCTCTGGGAGGCTCTGTGCTGCTCCAGCCAGTCAGCCGCCACTGACT 7408
OY 21 ----- 21
DB 7409 GGGCTGAAGTGTAGGAACGTGACCGTGGTCCAGCCAGCTCCACCTGACTGTGT 7468
OY 22 ----LeuProArgValSerProArgPheLeuAlaArgAlaPheHisProProArgValSerAla 40
DB 7469 CCCCTCCCAAGTg-AGCCCCCGCTTCTCCGAGCGTGGCAACCCCTCCCGTCTGACGC 7527

```

```

OY 41 ArgMetProThr----- 44
DB 7528 AGGATGCCACGCGTAGTGTGTCTAGCCTGCTTCTGCTCCCGGGTCTTCTTCAAC 7587
OY 44 ----- 44
DB 7588 TCCGTCCACACCCGAGGAGGCGCTTCTGTAAGGGCTGATCTGCCCCCTACCCCTTC 7647
OY 44 ----- 44
DB 7648 TGTGCGCTCGGCTGCGCTTCTCAGTCATCTTGGCTTCTGACCTTGTGCCAGGC 7707
OY 44 ----- 44
DB 7708 GCTCTCTTGTGACTAGCTAGATTTCTCTGAGACACCCACCTCCCTTCTTCCCCC 7767
OY 44 ----- 44
DB 7768 GATGTCGCGTCTGTTTCTGCGCTGCGCTCTTTCANCTATACCTTTCGCC 7827
OY 44 ----- 44
DB 7828 TCTGTCTCTTTCCTCCCTCCGCTGCTGCTCCCACTCCGATGTCCTCCCTTT 7887
OY 45 -----ArgArgTrpAlaProGlyThrGlyHisIlePheIly 56
DB 7888 GCCCTCTCCCGGCGGCTCTGCAAGAGGCGCTGGGCGCCAGCCAGTGTATACCAA 7947
OY 56 sCyGluHisThrArgProLysProGlyGluLeuAlaPheArgIlyGlyAspValIleThr 76
DB 7948 ATGGAGAGACACCGCCCAAGGACGAGGAGCTGCGCTTCCGAAGGGCGAGCGTGCAC 8007
OY 76 rIleLeuGluAlaGlyGlu----- 82
DB 8008 CATCTGAGGCGCTGCCA-GGTAGAGGTGCGCGGCGGTGTGCTTGGGGCTGGGG 8066
OY 83 -----As 83
DB 8067 GCTCCACGAGACACCCACCCACACCCACACTAACCCCTGCTTCTCCCTGGCAGAA 8126
OY 83 nLysSerTrpTyrArgValIlyHisIleThrSerGlyGlnGluLeuAlaIle 103
DB 8127 CAAGAGCTGGTACGCGTCAAGACACACACAGTGAAGAGGGGCTGTGGCAGCTGG 8186
OY 103 yAlaLeuArgIluArgIluAlaLeuSerAlaAspProLysLeuSerLeuMetPro----- 121
DB 8187 GCGCTCGGAGGCGGAGGCGCTCTCCGAGACCCCAAGCTCAGCTCAGCTGAGT 8246
OY 121 ----- 121
DB 8247 GGGGAGNACAGGGGCTGGGTAGGGGACAGCAAGTACCCCTCCACAGCCAGCTCT 8306
OY 122 -----TrpPheHisGlyLysIleSerGlyGlnGluAlaIle 133
DB 8307 GACCCACCCCTTCCGTGGCGCCAGAGTGTTCACAGGGAATCTCGGGCGAGAGGCTGT 8366
OY 133 lGlnGlnLeuGlnProProGluAspGlyLeuPheLeuValArgIleuSerAlaArgHisP 153
DB 8367 CCAGCAGCTCGAGCTCCCGAGATGGGCTGTCTGCTGGGAGTCCGGGCCACACC 8426
OY 153 oGlyAspTyrValLeuGlyValSerPheGlyArgAspValIleHisTyrArgValIleuH 173
DB 8427 CGGCGACTAGCTCTGTGCGCTGAGACTTTGGCGGAGCTCATTCACCTACCGCTGCTGCA 8486
OY 173 sArgAspGlyHisLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspMetVa 193
DB 8487 CCGGAGGCGGCACTCAATCATGATGAGCGGTCTTCTTGCACACCTCATGACATGTGT 8546
OY 193 lGlu----- 194
DB 8547 GGAAGTCTGTGCAACCCAGAGGCGCCACCCCGTCCCTGCACTGAGGGGAGCCAGGAGCTCC 8606

```

QY 194 ----- 194
Db 8607 GGAGACTCCCTACGTAGAGATAGAGGGGGGTACCTGACCTTCATCCTTCCAGC 8666
QY 194 ----- 194
Db 8667 AGCTGGGCGACACCATCCCTACCTCTGTGTGATGACCCGGGACCCTTGAGCTGTGT 8726
QY 194 ----- 194
Db 8727 GCTTCCCAATGTGAGATGAAGGGGTATTCCCAAGGGCCCCCAGGACCCCAAGGCC 8786
QY 194 ----- 194
Db 8787 CCCATCCCATCTCAGATGGGCGACAGAGAGACCCCAAGTAGATTCCCTTAATGCA 8846
QY 194 ----- 194
Db 8847 GAAATAGGGGGGTGATCCCTATGAGACTGCCCAATCTACAGGTGGCTGAGTACGCTGG 8906
QY 194 ----- 194
Db 8907 TTCTACCTCAGGAGACTAGAGTACTCTGTGGGAGACCCCAAGACCCCAAGTTGAT 8966
QY 194 ----- 194
Db 8967 ACCAATCCAGAGACTCCCCCTCTTTTGGCTTGCCCTTCGTTCCATGAAACC 9026
QY 194 ----- 194
Db 9027 AGCCTTCTCTCTCTCCCGGTCCACCCACCCACAGAGGCCACCCGGGAGCGAAC 9086
QY 195 -----H1 195
Db 9087 AGGATCTGGGGTTCCTCCCTGGGGGTGGGCTCATGGCTGCCACCATCTGCAGCA 9146
QY 195 -----H1 195
Db 195 sTyrSerIysAspIysGlyAlaIleCysThrIysLeuValArgProIysArgIysHisG1 215
Db 9147 TTACACCAAGACAAAGGGGCTATCTGCACCAAGCTGTGAGACCAAGCGGAACACGG 9206
QY 215 yThrIysSerAlaGluGluGluLeuAlaArgAla----- 226
Db 9207 GACCAAGTGGCGCGAGGAGAGAGCTGGCCAGGGG-TAGGGGAGGCCCAAGAGGCAAGACC 9265
QY 226 ----- 226
Db 9266 CCTTCCCTACCAAGCTTACCCAGTCCGGGAAGAGAGGCCCTGGGGCCCCGACCTTG 9325
QY 226 ----- 226
Db 9326 AGGCAGATCACCAAGCTGTCTCCATGCCAGCTATGGGATGAAGACCTGGGCTGCC 9385
QY 226 ----- 226
Db 9386 TGGGGGTGCCACAGACACAGATCAGTCTTATTCAATAGTCCCGACGTTCCGCACT 9445
QY 226 ----- 226
Db 9446 CATCTGTACACATCATCCACCATCTGAGCCAGAAACCTGGCCCTGAACCTCCACCAACC 9505
QY 226 ----- 226
Db 9506 CATCTCCAGCAATTCCTCTTGGCGAAGTTCCCAAGACTGGTGAATCTTCTCCCTCC 9565
QY 226 ----- 226
Db 9566 CCTCCCAAGCCAGCCTCAGGCCCTCCCATCAGACCGCTCTCTTCCAGAGGGGAGCT 9625
QY 226 ----- 226
Db 9626 CCCCTTCCAGCCAGAGCCTTGAAGTACAGCTCAAGGTTATCTTCTATCATGAACAAT 9685
QY 226 ----- 226

Db 9686 GCCTCCAGGAGAGAGAGGCCAGAGCTCCACACACTTCCAGAACCTGCCCTGATGGCCTG 9745
QY 226 ----- 226
Db 9746 AGCCACCTTCCAGCCTCATCTGCAAAAGCCCTTCTCCAGACATCCACCCCTCCAGGT 9805
QY 226 ----- 226
Db 9806 CAACGTGATCATGCTTTCTTCTCATGCTCTGAGCCATGTATATTCTGTCCCTTACCC 9865
QY 226 ----- 226
Db 9866 TGAATGCTTTTCCAGCCCTTAAGTAGTAATCTCAATATTAAAGTCAAAAGCAAT 9925
QY 226 ----- 226
Db 9926 GCCTCCTGACCCCATGCTGTCTCTCTGAAAGTATCTCTGCTGGGAGTCTGTAC 9985
QY 226 ----- 226
Db 9986 CCATCAGGATGAGCATCCAGGTGAGACTAGCATCACCTGGCAATGATCTCTTAAGAC 10045
QY 226 ----- 226
Db 10046 CGTGCATTTTGGAGGGGCACTTTCTGAGAGGAGAAGACGCTGCAGAACCGACAG 10105
QY 226 ----- 226
Db 10106 AGGGCAGTGGAGGGGACACACAGAGAGTCTTGGAGAAAGGTACAGGAGGCTAGAGCCC 10165
QY 226 ----- 226
Db 10166 TGAATATTCTGSAACCTTATGTGCTCATGTGTGTTTTTCTCTGTGTGTGTT 10225
QY 226 ----- 226
Db 10226 CTGGGGTGTGACAGAGTAAAGAACTGTAGAGAGAGCCAGAGGATTTTGGAGGAGC 10285
QY 226 ----- 226
Db 10286 AAGAGACCAAGAGAAAGGAGGAGCTTATCTGTGAGACATGAAGTGTTCCTGGCAC 10345
QY 226 ----- 226
Db 10346 TACCCCTTGGGCTGACTGGGAGTCTCTGGAGGGCAAGTGTGGGTCTGAGGCTCAGC 10405
QY 226 ----- 226
Db 10406 CAGGACAGCCCCAGCTTGAATATTCAACACACTATCCCTCTGAAGCCCTTTTGT 10465
QY 226 ----- 226
Db 10466 TCAGTCAACCTGAGATAGTCTCTGTGTTGCAACAACAATAATTGTAACAGAGATGGG 10525
QY 226 ----- 226
Db 10526 GTGTCTGATATATGCTTAGGGGCAATGCCAGGCCAGAGCCACTCAATGTGTATCCA 10585
QY 226 ----- 226
Db 10586 CAGCTGCAGACTGTGAGACATCAGCTAGTCTCTCACCAGTAGTAGAACCCCAATGGGA 10645
QY 227 -----GlyTrpLeu 230
Db 10646 GGATTTAGGGGGCCATTACCCCTGAGCTGGGTCTTCACTCCACAGCGGGCTGTATCT 10705
QY 230 uAsnLeuGlnHisLeuThrLeuGlyAlaGlnIleGlyGluPheGlyAla----- 248
Db 10706 GAACCTGACAGCAATTTGACATTTGGAGCACAGATCGGAGAGGAGATTTGGAGG-TGAGCT 10764
QY 248 ----- 248


```

OY 333 r1eu----- 334
Db 12984 TCTGTAGTGGAGACTCTCAGAGGTGTCTGGGACTGGGGTCTGGGGGCGAGATTCGAAG 13043
OY 334 ----- 334
Db 13044 ATCCATCAGAGGAAACTGAGGACGCGGAGGAGGCTTTGTGAGCTGTGTGCCCCAGG 13103
OY 334 ----- 334
Db 13104 CACCCGAGCTTCTGAGCCCTGATGTGGGATCTCTGGACCTGAGCCGCCACTGCCCC 13163
OY 335 -----HisValAlaGluGlyMetGluTyrLeuGluSerLysLysLeuValH 350
Db 13164 CTACTACCCCGACGCTGGCCGAGGAGCTGAGTACCTGGAGACGACAAAGCTTGTGC 13223
OY 350 1sArGspLeuAlaIaArgAsnIleLeuValSerLysLysLeuValAlaLysValSera 370
Db 13224 ACCGCACTGGCGCGCCGCAACATCTGTCTCAGAGGACCTGTGGCCAAAGTCAAGC 13283
OY 370 spPheGlyLeuAlaLysAlaGluArgLysGlyLeuAspSerSerArgLeuProValLysT 390
Db 13284 ACTTGGCTGGCCAAAGCGAGCGAGGAGGCTAGACTCAACCGCTCCGCTCAAGT 13343
OY 390 rPThrAlaProGluAlaLeuLysHisGly----- 399
Db 13344 GAGAGGCGCGCGAGGCTCTCAACACGCGGTGAGCCCTGCTCACATACCCCTGGGGCT 13403
OY 399 ----- 399
Db 13404 TTGGGTCGCCCGAGCTCTGCTGTATGACCCTGGGTATGCTCCTTGGCTCTCTGAAC 13463
OY 399 ----- 399
Db 13464 CCAAGGCAATGGCTATGCTCTCCCAAGAAAGCTCTGGCCCATATTTCTCTGTAGCC 13523
OY 399 ----- 399
Db 13524 CCCCCTGGGCTCAGTTTCTCCAGCTCTGMAAAAGCGTGGGCTCAGACTGAAAGGAA 13583
OY 399 ----- 399
Db 13584 GAATACCAAGCTCTCTCATGCACTAAGGCTGAGCTAGACTCACTGCGGGCGGCTTG 13643
OY 399 ----- 399
Db 13644 GGGGAGACTCAATTAGAGAGAACAGAGAGTGAAGGTCAACCCAGAGGCTTCTG 13703
OY 399 ----- 399
Db 13704 AGGAGCAGGGGCTGAGCAGAAACCTGAGGGGCTCTCTCCACCCCGCTCGGGGCC 13763
OY 400 -----LysPheThrSerLysSerAspValTyrSerPheGlyValLeuLeuTrpGluValP 418
Db 13764 CACAGAAAGTTCCACACAGAGTGGAGTGTCTGGAGTTTGGGGTGTCTGGAGAGTCT 13823
OY 418 heSerTyrGlyArgAlaProTyrProLysMetSer----- 429
Db 13824 TCTCATATGAGCGGGCTCCGTACCTAAATAGT-GAGCGGGGTCCAGAGGAGCACTG 13882
OY 429 ----- 429
Db 13883 GGTTCGGGCGAGTCCAGAGGCTGTGGCCCTGACCCCTGCTGTCAGCAGT 13942
OY 430 --LeuTysGluValSerGluAlaValGluLysGlyTyrArgMetGluProGluGlyC 449
Db 13943 CACTGAAGAAGGTGTCTGGAGCGCTGAGAAAGGGGTACCGCATGGAACCCCGAGAGGCT 14002
OY 449 ySProGlyProValHisValLeuMetSerSerCysTrpGluAlaGluProAlaArgArgP 469
Db 14003 GTCCAGGCGCCCGTGCAGCTCTCATGAGCAGTGTGTGGAGGAGAGCGCCGCGCGGG 14062
OY 469 roPheThrArgLysLeuAlaGluLysLeuAlaArgGluLeuArgSerAlaGlyAlaProA 489

```

```

Db 14063 CACCTTCGCGCAAACTGCGCGAGAGAGCTGCCCCGGGAGGCTACCCAGTGCAGTGCAG 14122
OY 489 1aSerValSerGlyGlnAspAlaAspGlySerThrSerProArgSerGlnPro 507
Db 14123 CCTCCGCTCAGGGGCGAGAGCGCCGAGGCTCCACTGTGCCCCGCAAGCCAGAGGCC 14178

RESULT 8
AAZ29701
ID AAZ29701 standard; cDNA; 2187 BP.
AC AAZ29701;
XX 22-MAR-2000 (first entry)
DE Wild-type human c-Src tyrosine kinase cDNA.
KW Angiogenesis; tyrosine kinase; Src; inhibition; activation; modulate;
KW human; viral expression vector; replication competent; mutant Src;
KW inflammatory disease; arthritis; rheumatoid arthritis; restenosis;
KW diabetic retinopathy; osteoporosis; cancer; ss.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 134..1486
FT /tag= a
FT /product= "Human c-Src tyrosine kinase"
FT /note= "Src used to modulate angiogenesis"
XX
XX W09961590-A1.
XX 02-DEC-1999.
XX
XX 28-MAY-1999; 99MO-US11780.
XX
XX 29-MAY-1998; 98US-0087220.
XX
XX (SCRI ) SCRIPPS RES INST.
XX
XX Cheresch DA, Ellicelri B, Schwartzberg PL;
XX WPI; 2000-116335/10.
XX P-P-SDB; AAY44448.
XX
XX Using tyrosine kinase Src for modulating angiogenesis in tissues useful
XX in, e.g. treatment of chronic articular rheumatism -
XX
XX Claim 1; Fig 3; 80pp; English.
XX
XX The present sequence is the cDNA, encoding the wild-type human c-Src
XX tyrosine kinase. This sequence encoding the Src protein can be used to
XX modulate angiogenesis. When the Src protein is inactivated, angiogenesis
XX is inhibited, while when it is activated, angiogenesis is potentiated.
XX The modified or variant Src can be used to treat inflammatory diseases
XX like, arthritis, rheumatoid arthritis, diabetic retinopathy, restenosis,
XX osteoporosis and cancer associated disorders.
XX
XX Sequence 2187 BP; 460 A; 651 C; 648 G; 428 T; 0 other;

Alignment Scores:
Pred. No.: 1.16e-87 Length: 2187
Score: 1247.50 Matches: 238
Percent Similarity: 71.59% Conservative: 82
Best Local Similarity: 53.24% Mismatches: 118
Query Match: 46.71% Indels: 9
DB: 21 Gaps: 3

US-09-977-260-2 (1-507) x AAZ29701 (1-2187)
OY 40 AlaArgMetProThrArg-----ArgTrpAlaProGlyThrGlnCys 53
Db 116 GCCAGAGCTCCTGAGAAAGTTCAGCAATACAGGCGCGCTGGCCATCCGCTACAGAAATGT 175

```

OY	54	lIerhryscysGluHstHstHrArgProLysProGluGluLeuAlaPheArgLysGly	73
Db	176	ATTCCAGATTCACAACTTCCACGGACATCCGACGACGAGCCCTGGCCATTAAGGAAC	235
OY	74	ValValThrIleLeuGluAlaCysGluLysSerTrpArgValLysHstHstHr	93
Db	236	GTGCTCAACCATGTGGCGGTCCACCAAGGACCCCAACTGTCAACAAAGCAAAACAAGTG	295
OY	94	SerLysGlnGluGlyLeuLeuAlaAlaCysAlaLeuArgGluArgGluAlaLeuSerLa	113
Db	296	--GGCGGTGAGGGCATCATCCACGCAATACATCCAGAAAGCGGAGCGGTGAAGCG	352
OY	114	AspProLysLeuSerLeuMetProTrpPheHstGlyLysIleSerGlyGluAlaVal	133
Db	353	GGTACCAAAAGTCAGCTCATGCTCTTGGTTCACGGCAAGATCACAGGGACAGCTCAG	412
OY	134	GlnGlnLeuGlnProProGluAspGlyLeuPheLeuValArgGluSerAlaArgHstPro	153
Db	413	CGGCTTCTGTATCCCGCCGAGACAGGCCCTGTCTGGCGCGGAGACCAACATACCCC	472
OY	154	GlyAspTrpValLeuCysValSerPheGlyLysArgAspValIleHstArgValLeuHst	173
Db	473	GGAACTACACACGCTGTGCTGCTAGCTGCGACGGCAAGGTGAGACACTACCGCATCTGAC	532
OY	174	ArgAspGlyHstLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspMetVal	193
Db	533	CATCCCAAGACATCAGCATCGACGAGAGAGGTGTACTTGGAAACCTCATGACGTGCTG	592
OY	194	GluHstArgSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLysArgLys	213
Db	593	GAGCACTACACTCCAGACAGATGAGATGTGTACGCCCTCATTTAAACCAAGGTCTATG	652
OY	214	HstGlyThrLysSerAlaGluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGln	233
Db	653	GAGGGACACAGTGGGGGCCAGAGATGATTCACGCCACGGCTGGGCCCTGAACATGAAG	712
OY	234	HstLeuThrLeuGlyAlaGlnIleGlyGluGlyLupheGlyAlaValLeuGlnGlyLys	253
Db	713	GAGCTGAAGCTGCTGCACACCATCATCGGAGGGGGGATTCGGAGACGTATGCTGGCGCAT	772
OY	254	TyrLeuGlyGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGluAlaPhe	273
Db	773	TACCGAGGGAACAAAGTCGCCGTCAAGTGAATTAAAGACAGCCACCTGCCCCAGGCTTC	832
OY	274	LeuAspGluThrAlaValMetThrLysMetGlnHstGluAsnLeuValArgLeuGly	293
Db	833	CTGCGTCAAGCCTCAGTCATGACGCAATCGCGGCATACCAACTGTGTCAGTCTCTGGCC	892
OY	294	ValIleLeuHstGln-----GlyLeuTyrIleValIleMetGluHstValSerLysGlyLys	311
Db	893	GTGATCTGGGAGGAGAAAGGCGGCGCTCAACTGCTCACTGAGTGAATGGCCAAAGGGGAGC	952
OY	312	LeuValAsnPheLeuArgTrpArgGlyArgAlaLeuValAsnThrAlaGlnIleLeuGln	331
Db	953	CTGTGGACATCACTCGCGGCTCAAGGGGTGGGTCAAGTGCCTGGGGCGAGACATGTCTCTCAAG	1012
OY	332	PheSerLeuHstValAlaGluGlyMetGluTyrLeuGlnSerLysLysLeuValHstArg	351
Db	1013	TTTCTCGCTAGATGTCTGCGAGGCCATGAAATACCTGGAGGGCAACAAATTTGCTGATCGA	1072
OY	352	AspLeuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaLysValSerAspPhe	371
Db	1073	GACCTGGCTGCCCAATGTCTGTGTCTTAAGGACAACTGGGCCCAAGGTCAAGCGACTTT	1133
OY	372	GlyLeuAlaLysAlaGluArgLysGlyLeuAspSerSerArgLeuProValLysTrpThr	391
Db	1133	GGTGTCAACAAAGAGGGCGTCCAGACACCAAGGACAGGGCAAGTGCCTCAAGTGGAGACA	1192
OY	392	AlaProGluAlaLeuLysHstGlyLysPheThrSerLysSerAspValTrpSerPheGly	411
Db	1193	GGCCCTGAGGCGCTGAGAGAGAAATTTCTCAATAGTGTGACGTGTGAGATTTCGGA	1255

OY	412	ValleuLeuTPGJLValaPheSeTyrGlyrArgAlaProTyrProLysMetSerLeuLys	1253	ATCCTTCCTCGGGAATATCTACTCCTTTTGCGGAGTGGCTTATCCAGAATTCCTCCCTGAAG	1
OY	432	GLuValSerGluAlaValaValaGLuLysGlyTyrArgMetGluProProGluGlyCysaProGly	1313	GACGTCGTCGCTCGGGGTGGAGACAGGGGCTACAAAGATGGATGGCCCGGAGCGCTGGCCGCC	1
OY	452	ProValHisValLeuMetSerSerCysTrrpGluAlaGluProAlaArgArgProProPhe	1373	GCAGCTCATGTGAAGCTCATGAAACACTGCGTCGACCTCGAGCAGCCGCGCATCGGCGCTTC	1
OY	472	ArgLysLeuAlaGluLysLeu	1433	CTACAGCTCCGAGACACTT	1453
DB	AAH28358	standard; cDNA; 2187 BP.			
AC	AAH28358;				
DT	05-SEP-2001	(first entry)			
DE		Nucleotide sequence of human tyrosine kinase protein Src.			
KM		Vascular permeability; tyrosine kinase protein; Src; Yes; stroke;			
KM		myocardial infarction; restenosis; trauma; blood vessel; atherosclerosis;			
KM		diabetic retinopathy; inflammatory disease; infection; arthritis;			
KM		adult respiratory distress syndrome; ARDS; rheumatoid arthritis;			
KM		diabetic retinopathy; psoriasis; neovascular glaucoma;			
KM		capillary proliferation; osteoporosis; cancer; ss.			
OS		Homo sapiens.			
FH	Key	Location/Qualifiers			
FT	CDS	134..1486			
FT		/*tag= a			
FT		/product= "Src"			
PN	WO200145751-A1.				
PD	28-JUN-2001.				
PF	22-DEC-2000; 2000WO-US35396.				
PR	22-DEC-1999; 99US-0470881.				
PR	29-MAR-2000; 2000US-0538248.				
PA	(SCRI) SCRIPPS RES INST.				
PI	Cheresh DA, Ellceiri B, Paul R;				
DR	WPI: 2001-417982/44.				
DR	P-PSDB; AAB84662.				
PT		Modulating vascular permeability in tissues, including inflamed tissue,			
PT		tissues associated with stroke, myocardial infarction, by contacting			
PT		the tissue with tyrosine kinase protein Src, Yes or their modified			
XX		forms			
XX					
PS	Disclosure; Fig 3; 133pp; English.				
CC		The specification describes a method for modulating vascular			
CC		permeability in a tissue suffering from a disease condition. The method			
CC		comprises contacting the tissue with a pharmaceutical composition			
CC		comprising tyrosine kinase protein Src, Yes or their mixtures or			
CC		nucleic acid expressing them. The method is useful for modulating			
CC		vascular permeability in tissues, including inflamed tissue, tissues			
CC		associated with stroke, myocardial infarction or other blockage of			
CC		normal flow, tissues undergoing restenosis, psoriatic, retinal tissue			
CC		and similar tissues. Pathologies which may be treated include			
CC		trauma to blood vessels, and other systemic pathological events such as			

CC atherosclerosis, diabetic retinopathy, inflammatory disease due to
 CC infection by microbial agents and arthritis. Other diseases which can
 CC be treated include adult respiratory distress syndrome (ARDS), rheumatoid
 CC arthritis, diabetic retinopathy, psoriasis, neovascular glaucoma,
 CC capillary proliferation in atherosclerotic plaques and osteoporosis and
 CC cancer associated disorders such as solid tumours, solid tumour
 CC metastases, angiodysplasias and hemangiomas. The present sequence
 CC encodes human Src, and is used in the method of the invention.
 XX

Sequence 2187 BP: 460 A; 651 C; 648 G; 428 T; 0 other:

Alignment Scores:

Score:	1,16e-87	Length:	2187
Best Local Similarity:	1247.50	Matches:	238
Percent Similarity:	71.59%	Conservative:	82
Best Local Similarity:	53.24%	Mismatches:	118
Query Match:	46.71%	Indels:	9
DB:	22	Gaps:	3

US-09-977-260-2 (1-507) x AAH28358 (1-2187)

QY 40 AlaArgMetProThrArg-----ArgTrpAlaProGlyThrGlnCys 53
 DB 116 GCCAGAGCTCTGAGAGATGTCAGCAATACAGGCCGCTGGCCATCGGTACAGAAATGT 175
 QY 54 IleThrLysCysGlnHisThrArgProLysProGlyGluLeuAlaPheArgLysGlyAsp 73
 DB 176 ATTCGCCAAGTACAACTTCCAGCGCACTGCCAGCAGAGACTGCTCTTCCAAAGAGAC 235
 QY 74 ValValThrIleLeuGlnAlaCysGlnAlaLysSerTrpTrpArgValHisIleThr 93
 DB 236 GTCTCACCATTGTGGCGCTGCACCAAGAGGCCCACTGTCACCAAGCAAAACAAAGTG 235
 QY 94 SerGlyGlnGluLeuLeuAlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAla 113
 DB 296 ---GGCGTGAAGGATCATCCAGCACTACCTGACAAAGCGGAGGCGTGAAGCG 352
 QY 114 AspProLysLeuSerLeuMetProTrpPheHisGlyLysIleSerGlyGlnGluAlaVal 133
 DB 353 GGTACCAACTGACCTCATGCTTGTTCACAGCAAGTACACAGGAGCAGGCTGAG 412
 QY 134 GlnGlnLeuGlnProProGlnAspGlyLeuPheLeuValArgLysSerAlaArgHisPro 153
 DB 413 CGGCTTGTGTACCGCGGAGACAGGCTGCTGCTGGGAGAGACCAACTACCC 472
 QY 154 GlyAspTrpValLeuLysValSerPheGlyArgAspValIleHisTrpArgValLeuHis 173
 DB 473 GGAGACTACACGCTGTGCTGAGCTCGACGCGCAAGGTGAGACATACCATCATGTAG 532
 QY 174 ArgAspGlyHisLeuThrIleAspGluAlaValPhePheCysAsnLeuMetVal 193
 DB 533 CATGCCAGCAAGCTCAGCATCGACGAGGAGGTGTACTTTGAGAACCTCATGCGTGTG 592
 QY 194 GlnHisTrpSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLysArgLys 213
 DB 593 GAGCACTACACCTCAACGACAGATGAGCTGTAGCGGCTCATTAACCAAGTCATG 652
 QY 214 HisGlyThrLysSerAlaGluGluAlaArgAlaGlyTrpLeuLeuAsnLeuGln 233
 DB 653 GAGGCGACAGTGGCGGCGAGAGTCTACCGGAGCGGCTGCGCTGAACATGAG 712
 QY 234 HisLeuThrLeuGlyAlaGlnIleGlyGluGlyGlnPheGlyAlaValLeuGlnGly 253
 DB 713 GAGCTGAAGCTCTGACAGCAATCGGAGAGGAGGAGTGTGAGACCTGATGCTGGCGAT 772
 QY 254 TyrLeuGlyGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGlnAlaPhe 273
 DB 773 TACCGAGGAGCAAGATGCGCTCAAGTCAATTAAAGACAGACCATCCAGGCTTC 832
 QY 274 LeuAspGluThrAlaValMetThrLysMetGlnHisGluAsnLeuValArgLeuGly 293
 DB 833 CTGGCTGAAGCCTCAGTATGACGCAACTGCGCATAGCAACTGCTGTGAGCTCGTGGGC 892

QY 294 ValIleLeuHisGln-----GlyLeuTrpIleValMetGlnHisValSerLysGlyAsn 311
 DB 893 GTGATGTTGGAGGAGAGAGGCGCTCATCTCATCTCATCTCATCTCATCTCATCTCATCT 952
 QY 312 LeuValAsnPheLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuGln 331
 DB 953 CTGTGACATACCTGCGGCTTAAAGGTCGTCAGTCTGCGGAGACTGCTCTCTCAAG 1012
 QY 332 PheSerLeuHisValAlaGluGlyMetGlyLeuGluLysSerLysLeuValHisArg 351
 DB 1013 TTCTCGCTAGATGCTCGAGAGGCTGAGCAATACCTGAGAGGCAACATTTCTGATCGA 1072
 QY 352 AspleuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaLysValSerAspPhe 371
 DB 1073 GACCTGCTGCTCCCGCAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1132
 QY 372 GlyLeuAlaLysAlaGluArgLysGlyLeuAspSerSerArgLeuProValLysTrpThr 391
 DB 1133 GGTCTCACCAAGAGGCGCTCCAGCACCAGCAGCGGCAAGCTGCAATGCAATGAGACA 1192
 QY 392 AlaProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspValTrpSerPheGly 411
 DB 1193 GCCCTGAGGCGCTGAGAGAGAAATTTCTCCATAGTCTGACGTGTGAGTTTGGA 1252
 QY 412 ValLeuLeuTrpGluValPheSerTrpGlyArgAlaProTrpProLysMetSerLeuLys 431
 DB 1253 ATCTTCTCTGGGAAATCTCTTGGGCGAGTGCCTTATCCAAAGATTCCTCCCTGAG 1312
 QY 432 GluValSerLysValAlaGluLysGlyTrpArgMetGluProProGluGluLysProGly 451
 DB 1313 GACCTGCTCCCTCGGCTGGAGAGGCTTACAAAGTGGATCCCCCGAGGCTCCGCC 1372
 QY 452 ProValHisValLeuMetSerCysTrpGluAlaGluProAlaArgArgProPhe 471
 DB 1373 GCAGTCTATAGATGATGAAGAACTGCTGACACCTGAGACCCGCAATGCGGCTCTTC 1432
 QY 472 ArgLysLeuAlaGluLysLeu 478
 DB 1433 CTACAGCTCCGAGAGCAGCTT 1453

RESULT 10
 ABK84602 standard; cDNA; 2187 BP.
 ID ABK84602:
 AC ABK84602:
 AC 14-AUG-2002 (first entry)
 DT Human cDNA differentially expressed in granulocytic cells #1173.
 XX
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; peritonitis; allergy;
 KW granulocyte activation; chronic inflammation; allergy.
 OS Homo sapiens.
 XX
 PN MO200228999-A2.
 XX
 PD 11-APR-2002.
 XX
 XX 03-OCT-2001; 2001MO-US30821.
 PF 03-OCT-2000; 2000US-237189P.
 PR (GENE-) GENE LOGIC INC.
 XX
 PA Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 PI
 XX

DR WPI: 2002-435328/46.
 XX Detecting granulocyte activation by detecting differential expression
 PT of genes associated with granulocyte activation, which serves as
 PT diagnostic markers that is useful for monitoring disease states and
 PT drug toxicity -
 XX Claim 1: SEQ ID NO 1173; 114bp; English.
 CC The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs). Identified by
 CC DNA chip analysis as given in the specification, and comparing
 CC the expression level to an expression level in an unactivated
 CC GC, where differential expression of Gs is indicative of GCA.
 CC Also included are modulating (M2) GA by contacting GC with an agent
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)
 CC for an agent capable of modulating GCA or an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease using the
 CC gene expression profile; (3) detecting (M4) an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease, by detecting the
 CC level of expression in a sample of the tissue of gene(s) from Gs, where
 CC the level of expression of the gene is indicative of inflammation;
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
 CC an allergic response in a subject, exposure of a subject to a pathogen
 CC or sterile inflammatory disease, by contacting a tissue having
 CC inflammation with an agent that modulates the expression of gene(s)
 CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
 CC modulating Gs; M3 is useful for screening an agent capable of modulating
 CC GCA preferably in an inflammation in a tissue; M4 is useful for
 CC detecting an inflammation (especially chronic) in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
 CC reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and M5 is useful for treating one of the above conditions. The present
 CC sequence represents a gene differentially expressed in granulocytes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 2187 BP; 460 A; 651 C; 648 G; 428 T; 0 other:
 Alignment Scores:
 Pred. No.: 1.16e-87 Length: 2187
 Score: 1247.50 Matches: 238
 Percent Similarity: 71.59% Conservative: 82
 Best Local Similarity: 53.24% Mismatches: 118
 Query Match: 46.71% Indels: 9
 DB: 24 Gaps: 3
 US-09-977-260-2 (1-507) x ABK84602 (1-2187)
 QY 40 AAlargmetProthArg-----ArGTrpAlarProGlyThGInCys 53
 DB 116 GCCAGACTCTCGAGAGATGTCAGCAATACAGCCGCGCATCGGACAGAAATGT 175
 QY 54 lIleThrLysGcSgUlnHstHrArgPrPolysProGlyGluLeuAlaPheArgLysGlyAsp 73
 DB 176 ATTCGCAAGTACCACTTCACAGGACATCCAGCAGGACCTGCTCTGCAAGGAGAC 235
 QY 74 ValValThrIleLeuGluAlaCysGluAsnLysSerTrpTrpArgValLysHisHisThr 93
 DB 236 GTGCTACACATTTGGCGCGTCACAGAGACCCCAACTGGTCAAAAGCAAAACAGGTG 295
 QY 94 SerLysGlnGluGlyLeuLeuAlaAlaGlyValAlaLeuArgGluArgGluAlaLeuSerAla 113
 DB 296 ---GGCGGTAGGAGCATATCCAGCCAACTACGTCAGAGAGCGGAGGCGTGAAGCG 352

QY 114 AspProLysLeuSerLeuMetProTrpPheHisGlyLysIleSerGlyGlnGluAlaVal 133
 DB 353 GGTACCAAACTACAGCTTATCCCTGTTGTCACAGGCAAGATACACAGGAGACAGCTAG 412
 QY 134 GlnGlnLeuGlnProProGluAspGlyLeuPheLeuValArgGluSerAlaArgHisPro 153
 DB 413 CGGCTCTGTACCCCGCGGAGACAGCCCTGTTCTGGTGGGAGAGACCAACTACCC 472
 QY 154 GlyAspTrpValIleCysValSerPheGlyArgAspValIleHisTrpArgValLeuHis 173
 DB 473 GGAACATACAGCGCTGTGTAGTGCAGCAGCAGGAGGAGAGACTACCGCATCATGAC 532
 QY 174 ArgAspGlyHisLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspMetVal 193
 DB 533 CATCCACAGACGCTACAGATGACAGGAGAGGTGTACTTGGAAACCTCATCAGCTGTG 592
 QY 194 GluHisTrpSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLysArgLys 213
 DB 593 GAGCATACACCTCAGCAGCAGATGAGACTGTACGCGCTCATTAACCAAGGTATG 652
 QY 214 HisGlyThrLysSerAlaGluGlnGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGln 233
 DB 653 GAGGACACAGTGGCGGCCAGATGATGTTACCCGACGCGCTGGCCCTCAACATGAG 712
 QY 234 HisLeuThrLeuGlyAlaGlnIleGlyGlnGluPheGlyAlaValLeuGlnGlyGlu 253
 DB 713 GAGCTGAAGCTCTGACAGACCATCGGAGAGGGAGTGGAGACGTGATCTGGCGAT 772
 QY 254 TyrLeuGlnGlyLysValAlaValAlaLysAsnIleLysCysAspValThrAlaGlnAlaPhe 273
 DB 773 TACGAGAGAACAAAGTGGCGGTCAAGTGATTAAGAACAGGACGACGCGCCAGGCTTC 832
 QY 274 LeuAspGluThrAlaValMetThrLysMetClnHisGluAsnLeuValArgLeuGly 293
 DB 833 CTGCTGAAGCTCATGATCATGACCACTGGGATGACCACTGGTGCACCTCTGGCG 892
 QY 294 ValIleLeuHisGln-----GlyLeuThrIleValMetGlnHisValSerLysGlyLys 311
 DB 893 GTGATCTGGAGAGAGAGAGGCGGCTTACTATCTGCTACTGTCATGATGCCAAGGAGGC 952
 QY 312 LeuValAsnPheLeuArgTrpArgGlyArgAlaLeuValAsnThrAlaGlnLeuLeuGln 331
 DB 953 CTGTGGACTACCTCGGGGTATAGGGGTGTCAGTGGTGGCGGAGACTGTCTCTCAG 1012
 QY 332 PheSerLeuHisValAlaGluGlyMetGlyLeuGlyLeuGlySerLysLeuValHisArg 351
 DB 1013 TTCTCGCTAGATGTCCTGGAGGCAATGAACTGGAGGCAACAAATTTGTCATCGA 1072
 QY 352 AspLeuAlaAlaArgAsnIleLeuValSerGlnAspLeuValAlaLysValSerAspPhe 371
 DB 1073 GACCTGGCTGCCCAATGTCTGTCTGTAGGACAAAGTGGCCCAAGGTACGACATT 1132
 QY 372 GlyLeuAlaLysAlaGluArgLysGlyLeuAspSerSerArgLeuProValLysTrpThr 391
 DB 1133 GGTCCTACACAGAGAGCGCTCCAGACCCAGAGGAGGCAAGCTGACGTCAGTGAAGTTCGGA 1192
 QY 392 AlaProGluAlaLeuLysHisGlyLysPheHisLeuSerLysSerAspValTrpSerPheLys 411
 DB 1193 GCCCTGAGGCCCTGAGAGAGAAATTCACATGATGTGACGTGTGAAGTTTCGGA 1252
 QY 412 ValLeuLeuTrpGluValPheSerTrpGlyArgAlaProTyrProLysMetSerLeuLys 431
 DB 1253 ATCCCTTCTGGAAATCTACTCTTTGGGAGAGTGCCTTATCCAAAGATTCCTGTAG 1312
 QY 432 GluValSerGluAlaValGluLysGlyTyrArgMetGluProProGlnGlyCysProGly 451
 DB 1313 GAGCTGTCGCTCGGCGGAGAGAGGCTACAAATGATGATGCCCGGAGCGCTGCCGCC 1372
 QY 452 ProValHisValIleMetSerCysTrpGluAlaGluProAlaArgArgProProPhe 471
 DB 1373 GCACTTATGAAGTATGAAGAACTGCTGGACCTGGAGCGCCCATGCGCCCTCTTC 1432

QY 472 ArgLyLeuAlaLulysLeu 478
 DB 1433 CTACAGCTCCGAGAGCAGCTT 1453
 RESULT 11
 ABL62918
 ID ABL62918 standard; DNA: 2187 BP.
 AC ABL62918;
 DT 15-MAY-2002 (first entry)
 DE Breast cancer related gene sequence SEQ ID NO:1255.
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytotoxic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 OS Homo sapiens.
 PN WC00194629-A2.
 PD 13-DEC-2001.
 PE 30-MAY-2001; 2001MO-US10838.
 XX 05-JUN-2000; 2000US-209473P.
 XX 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 (AVAL-) AVALON PHARM.
 PA Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 XX

PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.
 DR
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 XX Claim 1; SEQ ID 1255; 44pp; English.
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.
 XX
 XX Sequence 2187 BP; 460 A; 651 C; 648 G; 428 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 1.16e-87 Length: 2187
 Score: 1247.50 Matches: 238
 Percent Similarity: 71.59% Conservative: 82
 Best Local Similarity: 53.24% Mismatches: 118
 Query Match: 46.71% Indels: 9
 DB: 24 Gaps: 3
 US-09-977-260-2 (1-507) x ABL62918 (1-2187)
 QY 40 AlaArgMetProThrArg-----ArgTrpAlaProGlyThrGlnCys 53
 DB 116 GCCAGAGCTTCCTGAGAAGATGTAGCAATACAGCCGCCCTGCCATCCGCTACAGATGT 175
 QY 54 IleThrLysCysGlnHisThrArgProLysProGlyLeuAlaIleArgLysGlyAsp 73
 DB 176 ATGCCAAGTACCACTTCACAGCGACTGCCAGCAGCAGACCTGCTTCGCAAGAAGAGAC 235
 QY 74 ValValThrIleLeuAlaCysGlnAsnLysSerTrpTrpArgValLysHisIsthr 93
 DB 236 GTGCTGACCATTTGTGGCCGTACCAAGGACCCCACTGTGTACAAACCAAAAGGTG 295
 QY 94 SerGlyGlnGlyLeuLeuAlaIleGlyAlaLeuArgGlyLysLeuSerAla 113
 DB 296 ---GGCCGTGAGGATCATCCAGCCACACTACGTACAGAAGGGAGGCGTGAAGCG 352
 QY 114 AspProLysLeuSerLeuMetProTrpPheHisGlyLysIleSerGlyGlnGlyAlaVal 133
 DB 353 GGTACCAAACTCAGCTCATGCTGTTCACAGCAGCAATCATCACGGAGCAGCTGAG 412
 QY 134 GlnGlnLeuGlnProProGlnAspGlyLeuPheLeuValArgGlnSerAlaGlnHisPro 153
 DB 413 CGGCTTGTGTACCGCGGAGACAGAGCTGTCTCTGTGGGAGGACCAACTACCC 472
 QY 154 GlyAspTrpValLeuGlyValSerPheGlyArgAspValIleHisTrpArgValLeuHis 173
 DB 473 GGAGACTACAGCTGTGCTGACTCGACGAGGAGGTGAGAGCATACCGCATCATGTAC 532
 QY 174 ArgAspGlyHisLeuThrIleAspGlyAlaValPhePheCysAsnLeuMetAspVal 193
 DB 533 CATGCCAGCAGAGTACATCGACGAGGAGGTGTACTTTGAGAACTCATGTGAGCTGTG 592
 QY 194 GlnHisTrpSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLysArgLys 213

```

Db      593 GAGCACTACACCTGACGACGATGACTCTGTACGGCCCTCATTAACCAAGTCATG 652
Qy      214 HisGlyThrIysSerAlaGluGluLeuAlaArgAlaGlyTyrPleuLeuLeuGln 233
Db      653 GAGGCGACAGTGGCGCCAGAGATGTTCTACCGAGCGGTGGCCCTGACATGAG 712
Qy      234 HisLeuThrLeuGlyAlaGlnIleGlyGluGlyLeuPheGlyAlaValLeuGlnGlyGlu 253
Db      713 GAGCTGAAGCTGCTGCACACCATTCGGAGAGGGAGTTCGAGACGTCATCTGGCGAT 772
Qy      254 TyrLeuGlyGlnValAlaValAlaLysAsnIleLysCysAspValThrAlaGlnAlaPhe 273
Db      773 TACCGAGGAAACAAAGTCGCCGTCAATGATTAAGAAGACGCGACCTGCCAGCCCTTC 832
Qy      274 LeuAspGluThrAlaValAlaMetThrLysMetGlnHisGlnLeuValAlaGluLeuGly 293
Db      833 CTGGCTAAGCTAGTCATGACCAACTGGCGCATGCAACTGGTGCACCTCTCCGGC 892
Qy      294 ValIleLeuHisGln-----GlyLeuTyrIleValMetGlnHisValSerLysGlyAsn 311
Db      893 GTGATCTCGAGAGAGAGGCGGCTCTACTCTCTCATCTGATGATGATGCCAAGGGAGC 952
Qy      312 LeuValAsnPheLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuGln 331
Db      953 CTGTGTGACTACTGCGGTCTAGGGGTGCGTCACTGCTGGCGGAGACTGTCTCTCAAG 1012
Qy      332 PheSerLeuHisValAlaGluGlyMetGluTyrLeuGlnSerLysLysLeuValHisArg 351
Db      1013 TTCTCGTATGATGTCTCGAGGCGCATGTAATCTGAGGCAACAAATTTGTCATCGA 1072
Qy      352 AspleuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaLysValSerAspPhe 371
Db      1073 GACTGTGCTGGCCGACATGTCTGTCTGATGAGCAACAGCGCCAGCTAGGACTTT 1132
Qy      372 GlyLeuAlaLysAlaGluArgLysGlyLeuAspSerSerArgLeuProValLysTyrThr 391
Db      1133 GGTCTCACCAAGGAGGCGTCCAGCACCCAGCAGCGGCAAGCTGCAGTCAAGTGAACA 1192
Qy      392 AlaProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspValTyrPheGly 411
Db      1193 GCCCTTAGGCCCTGAGAGAGAGAAATTCCTCAATGCTGACGTGGAGTTGGGA 1252
Qy      412 ValLeuLeuTyrGluValPheSerTyrGlyArgAlaProLysPheLysMetSerLeuLys 431
Db      1253 ATCTTCTCTGGGAATCTACTCTTGGGGAGTGCCTTATCCAGAAATTCCTCTGAG 1312
Qy      432 GluValSerGluAlaValGluLysGlyTyrArgMetGluProProGluGlyCysProGly 451
Db      1313 GACGTCTCCCTCGGAGAGAGGCTACAAAGATGATGCCCGGAGCGCTCCGCC 1372
Qy      452 ProValHisValLeuMetSerSerCysTyrGluAlaGluProAlaArgProProPhe 471
Db      1373 GCATCTATGAAGTCATGAAGAACTGTGGCAGCTGAGCGCCGATGCGGCTCTCTC 1432
Qy      472 ArgLysLeuAlaGluLysLeu 478
Db      1433 CTACAGCTCCGAGAGACGCTT 1453

```

RESULT 12

AA246489 standard; DNA: 2420 BP.

AA246489;

13-MAR-2000 (first entry)

PKA substrate, Csk-family protein encoding DNA.

Protein kinase A; PKA; PKA signaling pathway; phosphorylation; cancer;
 kinase substrate; immunosuppressive disorder; proliferative disease;
 HIV infection; AIDS; immunodeficiency; autoimmune disease; Chk; Lsk; Hyl;
 systemic lupus erythematosus; Csk-family; Matk; Ctk; Bnk; Ntk; ss.

```

XX      OS      Homo sapiens.
XX      EH      Key
XX      FT      CDS
XX      FT      Location/Qualifiers
XX      PN      MO9962315-A2.
XX      PD      02-DEC-1999.
XX      PE      27-MAY-1999; 99WO-GB01680.
XX      PR      27-MAY-1998; 98NO-0002419.
XX      PR      30-DEC-1998; 98US-0114240.
XX      PA      (LAUR-) LAURAS AS.
XX      PA      (JONE/) JONES E L.
XX      PI      Hansson V, Levy FO, Mustelin T, Skalhogg BS, Sundvold V, Tasken K;
PI      Vang T, Altman A, Munshi A;
PI      WPI: 2000-086801/07.
PI      P-PSDB: AAY49418.
PT      Altering the activity of protein kinase signaling pathways, used for
PT      treating immunosuppressive disorders, e.g. AIDS, proliferative
PT      disorders, e.g. cancers or autoimmune diseases
PS      Claim 11: Page 87-89; 11pp; English.
XX      CC      The invention provides a novel method of altering the activity of the
XX      CC      protein kinase A (PKA) signaling pathway in a cell that comprises
XX      CC      altering the extent of phosphorylation of one or more PKA substrates, or
XX      CC      kinase substrates downstream in the PKA signaling pathway. Pharmaceutical
XX      CC      compositions containing a nucleic acid molecule that encodes a PKA
XX      CC      substrate, or fragment, precursor or functionally equivalent variant,
XX      CC      where the sequence is modified to alter its susceptibility to
XX      CC      phosphorylation by PKA can be used for treating a disorder exhibiting
XX      CC      abnormal PKA signaling activity, immunosuppressive disorders or
XX      CC      proliferative diseases. They can be used for treating e.g. HIV
XX      CC      infection, AIDS, common variable immunodeficiency or cancers. Conditions
XX      CC      in which upregulation of the PKA pathway is required, such as autoimmune
XX      CC      disease, e.g. systemic lupus erythematosus, may also be treated. The
XX      CC      present sequence represents a DNA sequence encoding a PKA substrate,
XX      CC      wherein the substrate is in the Csk-family, preferably Csk, Chk, Lsk,
XX      CC      Hyl, Matk, Ctk, Bnk or Ntk.
XX      SQ      Sequence 2420 BP; 471 A; 771 C; 703 G; 475 T; 0 other;
XX      Alignment Scores:
XX      Pred. No.: 1.32e-87 Length: 2420
XX      Score: 1247.50 Matches: 238
XX      Percent Similarity: 71.59% Conserv: 82
XX      Best Local Similarity: 53.24% Mismatches: 118
XX      Query Match: 46.71% Indels: 9
XX      DB: 21 Gaps: 3
XX      US-09-977-260-2 (1-507) x AA246489. (1-2420)
Qy      40 AlaArgMetProThrArg-----ArgTyrAlaProGlyThrGlnCys 53
Db      395 GCCAGAGCTCTGAGAGATGTCAGCAATACAGGCCCGCTGCGATCGAGCAATGT 454
Qy      54 IleThrLysCysGluHisThrArgProLysProGlyLeuLeuAlaPheArgLysGlyAsp 73
Db      455 ATTGCCAAGTACCACTTCCAGGCGACTGCGAGCAGACCTGCTTCTGCAAGAGAGAC 514
Qy      74 ValValThrIleLeuGluAlaCysGluAsnLysSerTyrArgArgValLysHisThr 93
Db      515 GTGCTCACATGTGGCGCGTCCACAGAGACCCCACTGTGTCAAGGCCAADAACAGGTG 574
Qy      94 serGlyGlnGlyLeuLeuAlaGlyAlaLeuArgGluArgGluAlaLeuSerAla 113

```



```

Db      575  ---GGCCGTGAGGGCATCATCCAGCCACTACGTCAGAGAGCGGGAGGGCGGTGAAGCGC 631
Qy      114  AspProLysLeuSerMetProTrrPheHisGlyLysIleSerGlyInGluAlaVal 133
Db      632  GGTACCAACTCAGCCTCATGCTGTGGTCCACGGCAAGATCAACAGGGAGAGCGCTGAG 691
Qy      134  GlnGlnLeuGlnProProGluAspGlyLeuPheLeuValArgGluSerAlaIleArgHisPro 153
Db      692  CGGCTTGTGACCCGCGGAGACAGCGCTGCTCTGGTGGGAGACACCAACTACCC 751
Qy      154  GlyAspTyrValLeuGlyValSerPheGlyArgAspValIleHisTyrArgValLeuHis 173
Db      752  GGAGACTACACGCTGCTGCTGAGCTGCGAGCGCAAGGTGAGAGCTACCCCATCATGTAC 811
Qy      174  ArgAspGlyHisSerPheIleAspGluAlaValPhePheCysAsnLeuMetAspMetVal 193
Db      812  CATGCCAGCAAGCTCAGCATCGAGAGAGAGGTGACTTGTAGAACTCATGACAGCTGTG 871
Qy      194  GlnHisTyrSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLysArgLys 213
Db      872  GACACATACACCTCAGACGAGATGAGACTCTGTACGCGCTCATTAACCAAGGTCATG 931
Qy      214  HisGlyThrLysSerAlaGlnGluGluLeuAlaArgAlaGlyLysPheLeuAsnLeuGln 233
Db      932  GAGGGCACAGTGGCGCCGAGATGAGTTTACCGGAGCGGCTGGCCCTGAGACATGAG 991
Qy      234  HisLeuThrLeuGlyAlaGlnIleGlyGluGlyIlePheGlyAlaValLeuGlnGlu 253
Db      992  GAGCTGAAAGCTGCTGCGAGACATCGGAGAGGGAGTGTGAGAGCTGATGCTGGCGAT 1051
Qy      254  TyrLeuGlyGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGlnAlaPhe 273
Db      1052  TACCGAGGGAACAAGTGGCGCTCAGTGTAGAACCAACCCACTGACCCAGCGCTTC 1111
Qy      274  LeuAspGluThrAlaValMetThrLysMetGlnHisGluAsnLeuValArgLeuLeuGly 293
Db      1112  CTGGCTGAAAGCTCAGTCATGACGACATCGGATGACCACTGGGAGCTGCTGGGC 1171
Qy      294  ValIleLeuHisGln-----GlyLeuTyrIleValMetGlnHisValSerLysGlyAsn 311
Db      1172  GTGATCGTGGAGAGAGAGCGGCTGTACATGCTGCTGATGATGATGATGATGATGATG 1231
Qy      312  LeuValAsnPheLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuGln 331
Db      1232  CTGTGTGACATCACTGCGGTCTAGCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1291
Qy      332  PheSerLeuHisValAlaGluGlyMetGluTyrLeuGlnSerLysLysLeuValHisArg 351
Db      1292  TTCTCGCTAGATGTCTGCGAGGCAATGATACCTGGAGGGCAACAATTTCTGTCATCGA 1351
Qy      352  AspLeuAlaIleArgAsnIleLeuValSerGluAspLeuValAlaLysValSerAspPhe 371
Db      1352  GACCTGGCTGCCCGCATGTGCTGTGAGAGACAGCTGGCCAAAGTCACAGCACTTT 1411
Qy      372  GlyLeuAlaLysAlaGluArgLysGlyLeuAspSerSerArgLeuProValLysTrpThr 391
Db      1412  GGTCTCACCAAGAGGCGCTCCACACCAGCAAGCGGCAAGTGTGCTGCTGCTGCTGCTG 1471
Qy      392  AlaProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspValTrpSerPheGly 411
Db      1472  GCGCCGAGGCGCTGAGAGAGAGAAATTTCTCACAATAGTCTACGTGTGAGATTTCCGA 1531
Qy      412  ValLeuLeuTrpGluValPheSerTyrGlyArgAlaProTyrProLysMetSerLeuLys 431
Db      1532  ATCTCTCTCTGGAAATCTACTCTTTGGCGAGTGCCTTATTCAGAAATTTCCCTGAG 1591
Qy      432  GluValSerGluAlaValGluLysGlyTyrArgMetGluProProGluGlyCysProGly 451
Db      1592  GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1651
Qy      452  ProValHisValLeuMetSerCysTrpGluAlaGluProAlaArgArgProProPhe 471

```

```

Db      1652  GCACTCTATAGTATGAAGAACTGCTGCGACCTGAGACCGCCGATGCGCCCTCCTTC 1711
Qy      472  ArgLysLeuAlaGluLysLeu 478
Db      1712  CTACAGCTCCGAGAGACAGCTT 1732

RESULT 13
ACC50120
ID      ACC50120 standard; cDNA; 2420 BP.
XX
AC      ACC50120;
XX
DT      12-JUN-2003 (first entry)
XX
DE      Breast cancer associated cDNA sequence SEQ ID NO:87.
XX
KW      Human; breast cancer; cytostatic; gene therapy; gene; ss.
XX
OS      Homo sapiens.
XX
PN      WO2003004989-A2.
XX
PD      16-JAN-2003.
XX
PF      21-JUN-2002; 2002WO-0519669.
XX
PR      21-JUN-2001; 2001US-299887P.
PR      27-JUN-2001; 2001US-301572P.
PR      18-JUL-2001; 2001US-306501P.
PR      25-SEP-2001; 2001US-325002P.
PR      05-MAR-2002; 2002US-362585P.
PR      14-MAY-2002; 2002US-380391P.
XX
PA      (MILL-) MILENIUM PHARM INC.
XX
PI      Lillie J, Gannavathu M, Ghatt K, Hoersch S, Kamatkar S, Mertens M;
PI      Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE, Bast RC;
PI      Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;
XX
DR      WPI; 2003-210381/20.
XX
PT      P-PSDB; ABR47428.
XX
PT      Breast cancer diagnosis or treatment by comparing the level of
PT      expression of a marker in a patient sample with that in the control
PT      non-breast cancer sample
XX
PS      Claim 1; SEQ ID 87; 128pp: English.
XX
CC      The present invention describes a method for assessing whether a patient
CC      is afflicted with breast cancer. The method comprises comparing the level
CC      of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
CC      ABR47386 to ABR47632) in a patient sample and the normal level of
CC      expression of the marker in a control non-breast cancer sample, where a
CC      significant increase in the level of expression of the marker in the
CC      patient sample and the normal level is an indication that the patient is
CC      afflicted with breast cancer. The breast cancer associated sequences
CC      from the present invention have cytostatic activities and can be used in
CC      gene therapy. The method is useful for diagnosing and treating breast
CC      cancer.
CC      N.B. The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from WIPO
CC      at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ      Sequence 2420 BP; 471 A; 771 C; 703 G; 475 T; 0 other;
XX

Alignment Scores:
Pred. No.:      1,32e-87      Length:      2420
Score:          1247.50      Matches:      238
Percent Similarity: 71.59%      Conservative: 82
Best Local Similarity: 53.24%      Mismatches: 118
Query Match:    46-71%      Indels:      9
DB:             25          Gaps:      3

```

US-09-977-260-2 (1-507) x ACC50120 (1-2420)

```

OY 40 AlaArgMetProThrArg-----ArgTrpAlaProGlyThrcIncs 53
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 395 GCCAAGAGCTCCGAGAAAGATGACGACATACAGCGCCCTGGCCATCCGATACAGAAATG 454
OY 54 IleThrLysCysGlnHisThrArgProLysProGlyLLeuAlaPheArgLysLysAsp 73
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 455 ATTGCCAAGTACAACTCCAGCGGCGCTCCGAGAGCACTGCCCTTCTGCAAGAGAAC 514
OY 74 ValValThrIleLeuGluAlaLysGlnLysSerTrpTyrArgValLysHisThr 93
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 515 GTGCTACCACTATGTGGCGGTACCAAGGACCCCACTGTACAAAGCAAAACCAAGGTG 574
OY 94 SerGlyGlnGluGlyLeuAlaLysGlnLysSerTrpTyrArgValLysHisThr 113
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 575 ---GGCGGTGAGGGGATATCCAGCCCACTACGTCGAGGAGGAGGCGGTGAGGCG 631
OY 114 AspProLysLeuSerLeuMetProTrpPheHisGlyLysLysGlyGlnGluAlaVal 133
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 632 GGTACCAAACTCAGCTCATGCTGTGGTCCAGCGCAAGATCACACGGGAGCAGGCTGAG 691
OY 134 GlnGlnLeuGlnProGlyLysAspGlyLeuPheLeuValArgGlnSerAlaArgHisPro 153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 692 CGGCTTCTGTACCCGCGGAGACAGCGCTGTCTGTGGTGGGAGAGCACCACATACCCC 751
OY 154 GlyAspTyrValLeuCysValSerPheGlyArgAspValIleHisTyrArgValLeuHis 173
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 752 GGAAGACTACAGCGTGTGGTGTAGCTGCGAGCGCAAGGAGGAGCACTACCGCATATGAC 811
OY 174 ArgAspGlyHisLeuThrIleAspGlnAlaValPhePheCysAsnLeuMetAspMetVal 193
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 812 CATCCAGCAAGCTCAGATGAGAGAGAGGTGACTTGTGGAACCTCATCATCAGCTGAGTG 871
OY 194 GlnHisTyrSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLysArgLys 213
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 872 GAGCAGTACAGCTCAGAGCGAGATGAGACTGTCTGACGCGCCCTCATTAACCAAGGTCATG 931
OY 214 HisGlyThrLysSerAlaGlnGluGlnLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGln 233
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 932 GAGGCGACAGTGGCGGCGCAGATGATGTACCGAGCGGCTGGCGCTGACATGAG 991
OY 234 HisLeuThrLeuGlnGlyAlaGlnIleGlyGlnGlyLysLysGlyAlaValLeuGlnGly 253
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 992 GAGCTGAGAGCTGTCGACAGCATCGGAGAGGGGAGGTGGAGACGATGCTGGCGAT 1051
OY 254 TyrLeuGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGlnAlaPhe 273
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1052 TACCGAGGAAACAAAGTCGCGCTCAGATGATTAAGAACAGCCACTGCCAGGCTTC 1111
OY 274 LeuAspGlnThrAlaValMetThrLysMetGlnHisGlnLysLeuValArgLeuGly 293
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1112 CTGGCTCAGAGCTCAGTACATGACCACTGCGGATGACCAAGCTGTCAGCTCTGGGC 1171
OY 294 ValIleLeuHisGln-----GlyLeuThrIleValMetGlnHisValSerLysGlyAsn 311
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1172 GTGATCTGGAGGAGAGAGGGGCGCTCTACATGCTCACTAGTACATGCGCAAGGGAGC 1231
OY 312 LeuValAsnPheLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuGln 331
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1232 CTGTGGAGTACACTGCGGTAGGGGTGGTCAAGTGTGGCGGAGACTGTCTCTCAAG 1291
OY 332 PheSerLeuHisValAlaGlnGlyMetGlnTyrLeuGlnLysSerLysLeuValHisArg 351
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1292 TTCCTGCTAGATGCTCGAGGCGCATGAAATACCTGAGGCAACAAATTCGTCATCGA 1351
OY 352 AspLeuAlaAlaArgAsnIleLeuValSerGlyAspLeuValAlaLysValSerAspPhe 371
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1352 GACTGTGCTGCCCGCATGTGCTGCTGTGAGGACACACTGGCCAAAGGTGACGACTTT 1411
OY 372 GlyLeuAlaLysAlaGlnArgLysGlyLeuAspSerSerArgLeuProValTyrTrpThr 391
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1412 GGTCTCACCAGAGAGGCTCCAGCACCCAGACAGCGGCAAGCTGCCAGTCAAGTGGACA 1471

```

```

OY 392 AlaProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspValTrpSerPheGly 411
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1472 GCCCTGAGGCCCTGGAGAGAGAGAAATTCACCTAAGTGTGACGTGTGGAGTTGCGGA 1531
OY 412 ValLeuLeuTrpGluValPheSerTyrGlyArgAlaProTyrProLysMetSerLeuLys 431
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1532 ATCCTTCTCTGGAAATATCTACTCTTTGGCGGAGTGCCTTATCCAGAAATTCCTCGAAG 1591
OY 432 GlnValSerGlnAlaValGlnLysGlyTyrArgMetGlnProProGlnGlyCysProGly 451
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1592 GACGTGCTGCTCGGTGAGAGAGGCTACAAAGATGATGCCCGGAGGCTGCCCGCC 1651
OY 452 ProValHisValLeuMetSerSerCysTrpGlnAlaGlnProAlaArgArgProPhe 471
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1652 GCAGTCTATGAAAGTATGAAAGAACTGCTGGCAGCTGGAGACCGCCATGCGGCTCCTTC 1711
OY 472 ArgLysLeuAlaGlnLysLeu 478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1712 CTACAGCTCCGAGAGCAGCTT 1732

RESULT 14
AAI93804
ID AAI93804 standard; cDNA; 2466 BP.
XX
AC AAI93804;
XX
DE 06-NOV-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 13864.
DE
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US04927.
XX
PF 28-FEB-2000; 2000US-0515126.
XX
PR 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI: 2001-514838/56.
XX
PT P-PSDB; AAO13873.
XX
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders.
XX
PS Claim 1; SEQ ID NO 13864; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO

```


xx Disclosure; Page 313-314; 389pp; English.

xx The invention describes a method of treating genetically-defined disease

cc associated with chromosomal aberrations yielding oncogenic fusion

cc proteins (I), treating cancerous cells containing (I) in a heterogeneous

cc cell population, treating proliferative diseases associated with mutant

cc protein or cellular protein isoforms (II) dependent on heat shock

cc protein (HSP)-90, or selectively treating cells expressing (II)

cc involving administering HSP90-inhibitor. The method is useful for

cc treating genetically-defined disease with chromosomal aberration yielding

cc oncogenic fusion protein, treating cancerous cells containing fusion

cc protein in heterogeneous cell population, treating proliferative disease

cc (e.g. rheumatoid arthritis or cancer) associated with mutant protein or

cc cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.

cc p53), or selectively treating cells expressing mutant protein or cellular

cc protein isoform in a patient heterozygous for (II). The method is useful

cc for treating a disease e.g. haematopoietic disorder such as T or B cell

cc lymphoma, chronic myeloid leukaemia (CML), APL, ALL, MLL, NHL and CML,

cc or a disease characterised by a solid tumour such as papillary thyroid

cc carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and

cc synovial sarcoma. The method is also useful for treating viral

cc infections. This sequence encodes a human oncogenic protein.

xx

xx Sequence 2442 BP: 474 A; 779 C; 708 G; 481 T; 0 other:

xx

Alignment Scores:

Pred. No.:	1,266-83	Length:	2442
Score:	1196.50	Matches:	234
Percent Similarity:	72.418	Conservative:	81
Best Local Similarity:	53.798	Mismatches:	116
Query Match:	44.808	Indels:	5
DB:	24	Gaps:	2

US-09-977-260-2 (1-507) x ABS73326 (1-2442)

QY 47 TTPALAPROGLYTHRGILGNCYSILETHRLYSCYSGIUNHISHTHARGPROLYSPROGLYGLU 66

DB 460 TGGCCATCCGGTACAGATGTATTGCAAGTACACTCCACGGGACGCGGAGAGGAC 519

QY 67 LEUALAPHEATRGVGLYASPVAYALVATHRIELEUGIUALCYSGIUNHISLYSERTTP 86

DB 520 CTGCCCTCTGCAAGGAGAGCGTGCTCACTGCGCGGCGGACCAAGGACCCCACTGG 579

QY 87 TYRARGVALYSHSHSTHRSERGLYNGIUNGLYLEUVALAALGLYALAEUARG 106

DB 580 TACAAAGCCAAAGAGAGTG---GGCCGTAGGAGCATCATGCCAGCCACTACGTCCAG 636

QY 107 GIUARGIUALAEUSERIALASPRPROLYSEUSERLEUMETPRG-TYRPHENISGLYLY 126

DB 637 AAGCGGAGGGGGGTAGAGGGGTACCAAACTCACCTCATGCGGTGAGTTCCAGGCA 696

QY 126 SILESERGLYNGIUALAVALGINGIUNGLINPROFROGLUASPGIYLEUVALA 146

DB 697 GATCACAGGAGGAGCGGTGAGCGGCTGTCTACCCGCGGAGACAGCGCTGTCTGAT 756

QY 146 IARGIUSERIALASPRPROGLYASPTVALLEUCYSVALSERPHEGLYARGSPYA 166

DB 757 GCGGAGAGGACCAACTACCCCGGAGACTACACGCTGTGCGTAGCTCGAGCGGCAAGT 816

QY 166 ILLEHISTYRARGVALLEUHSARGASPDYHISLEUTHRIEASPGIUALAVALAPHER 186

DB 817 GAGAGCATACCGCATCATGTACCATGCCAGCAAGCTACAGCATCGAGAGAGGTACTT 876

QY 186 ECYASALEUMETASPMETVALGINHISTYRSELYSASPLYGLYALALECYSTHLY 206

DB 877 TGAGAACCTCATGCGCTGTGAGACACTACACTCAGACGAGATGAGTCTGTACCGG 936

QY 206 SLEUVALARGPROLYSARGVLYSHISGLYTHRLYSESLAGIUNGLIUNGLUVALAARGAL 226

DB 937 CCTCTTTAACCAAGGTCATGAGGAGCAAGTGGCGCCAGGATGAGTTTACCGGAG 996

QY 226 AGLYTRPLEUVALAENGLINHISLEUTHRIEUGIYALAGIINLEGLYGLUGIUNPH 246

DB 997 CGCGTGGGCGCTGAACATGAGAGAGCTGAAGCTCTGCGAGACCAATCGGAGAGGGGAGT 1056

QY 246 EGIYALAVALEUENGLYGLUTYRLEUGIYGLINHISVALAVALYASANIILEYSCY 266

DB 1057 CGGAGACGTGATGCGTGGGCGATTCACCGGAGAACAAATCCCGTCAAGTCAATTAAGA 1116

QY 266 SASPVATHRIAGIUALAPHELEUASPGIUTHRIALAVALETHRLYSEMETGLINHISGL 286

DB 1117 CGAGGCCACTGCGCCAGGCTTCTTGCGTGAAGGCTCACTCATAGCGCAACTCGGCAATG 1176

QY 286 UASILEUVALARGLEUENGLYVALILEUENHISGLN-----GLYEURTYRILEVALME 304

DB 1177 CAACCTGTGACAGCTCTCGGCGGTGATGTGAGAGAGAAAGGCGGCTCATATCGTCAAC 1236

QY 304 TGLINHISVALSERLYSGIUNGLYASNPHELEUARGHTRARGIYALAGIUALAEUVA 324

DB 1237 TGAGTACATGGCCCAAGGGGAGCGCTTGCTGACTCTGCGGTAGGGGTGCGTCAAGTCT 1296

QY 324 IASNTHRALAGIUNGLINPHESERLYSHISVALAAGIUNGLYMETGLIUTYRLEUGI 344

DB 1297 GGGCGGAGACTGTCTCCATCAAGTTCTGCTAGATGTCTGCGAGGCGCATGGAATCTGGA 1356

QY 344 USERLYSILEUVALHISARGASPLEUVALAARGASNIILEUVALSERGLUASPLE 364

DB 1357 GGGCAACAATTTGCGCATCGAGACTGCTGCCCGCAATGTGCTGTCTGAGAGCA 1416

QY 364 UVALAVALYVALSERASPRPHEGLYLEUVALYSHLAGIUNRGVLYSGIYLEUASPSERSE 384

DB 1417 GGTGGCAAGGTGACGAGACTTTGGTCTCCACCAAGGAGCGCTCCACACC-CAAGACACGGG 1475

QY 384 RARGLEUPROVALYSTRPHTHALPROGLUALAEUVALYSHISGLYLYSPHETHSERLY 404

DB 1476 CAAGTGGCAGTCAAGTGGAGACAGCCCTGAGAGGCCCTGAGAGAGAAATTTCTCACATTA 1535

QY 404 SSERASPVATHRPSERPHEGLYVALLEUENUTRPGIUNVALPHESERTYRGLYARGALAPR 424

DB 1536 GTCTGACGTGTGAGATTTGCGAATCTCTCTGCGAATAATCTCTTGGGCGAGTCC 1595

QY 424 OTYRPROLYSMETSERILEUVALYVALSERGLUALAVALGILYSGIYTYRARGMETGL 444

DB 1596 TTATCCAAAGAAATCCCTGAGAGAGCTGCTCCCTGCGGTGAGAAAGGCTTACAAAGATGA 1655

QY 444 UPROFROGLIUCYSPROGLYPROVALHISVALLEUMETSERCYSTRPGIUALAGI 464

DB 1656 TGCCCCCAGCGGCTCGGCGGCGAGCTGTATGAATCATGAAGAACTGCTGCGACACTGGA 1715

QY 464 UPROVALARGARGPROFROPHETARGVLYSLEUVALAGIULYSLEU 478

DB 1716 CGCCGCAATGGCGGCTCTCTCTCATACGCTCCGAGACAGACTT 1758

Search completed: August 1, 2003, 20:02:54
Job time : 504 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2003, 18:16:35 ; Search time 6032 Seconds

(without alignments)
3438.525 Million cell updates/sec

Title: US-09-977-260-2
Perfect score: 2671
Sequence: 1 MAGRSLVSMRAFHCDSAE.....PASVSGDADGSPRSQEP 507

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2688711 segs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO/spool/US09977260/runat_29072003_092741_14619/app.query.fasta.1.647
-DB=genEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptlo -NHEP-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09977260.ecgn.1.1.5265.grunat_29072003_092741_14619 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2671	100.0	1942	6 AR053292	AR053292 Sequence
2	2671	100.0	1942	6 E10345	E10345 cDNA encodi
3	2671	100.0	1968	6 HSHYLTJK	X77278 H sapiens H
4	2671	100.0	1997	6 BD176710	BD176710 Method of
5	2671	100.0	2000	6 AR179662	AR179662 Sequence
6	2671	100.0	2000	6 AR274487	AR274487 Sequence
7	2671	100.0	2180	6 BC000114	BC000114 Homo sapi
8	2671	100.0	2180	6 BC003109	BC003109 Homo sapi
9	2662	99.7	1521	6 AR053291	AR053291 Sequence
10	2607.5	97.6	1744	6 S71669	S71669 1st-Leukocy
11	2591.5	97.0	1987	6 AR084686	AR084686 Sequence
12	2591.5	97.0	1987	6 AR204704	AR204704 Sequence
13	2591.5	97.0	1987	6 H0MMATK	L18974 Human tyros
14	2544	95.2	1804	9 HSM802388	AL13754 Homo sapi
15	2541.5	95.2	2073	6 AK055395	AK055395 Homo sapi
16	2444	91.5	1398	6 AR053290	AR053290 Sequence
17	2341	87.6	1734	10 MDSNTK	L27738 Mus musculu
18	2290.5	85.8	1911	10 MDSCTK	D45243 Mouse mRNA
19	2279.5	85.3	1838	10 RATABTK	L34542 Rattus norv
20	2277.5	85.3	1656	10 S77473	S77473 VNK-nontre
21	2274.5	85.2	1651	10 MMU05210	U05210 Mus musculu
22	2020.5	75.6	1713	6 AX644057	AX644057 Sequence
23	1512	60.4	246796	2 AC094643	AC094643 Rattus no
24	1592.5	59.6	247196	2 AC073822	AC073822 Mus muscu
25	1591.5	59.6	218817	2 AC114704	AC114704 Rattus no
26	1540.5	57.7	8212	6 AX306276	AX306276 Sequence
27	1540.5	57.7	8212	10 MMHYLTJK	X83972 M. musculus
28	1501	56.2	43190	9 AC005777	AC005777 Homo sapi
29	1469.5	55.0	16389	6 AX644059	AX644059 Sequence
30	1269	47.5	738	6 AR053289	AR053289 Sequence
31	1253.5	46.9	1353	5 CHRSRCKA	M85039 Chicken src
32	1252	46.9	2292	10 BC018394	BC018394 Mus muscu
33	1252	46.9	2363	10 BC052006	BC052006 Mus muscu
34	1247.5	46.7	2187	6 AX330746	AX330746 Sequence
35	1247.5	46.7	2187	9 HSCSRCKIN	X59932 Human mRNA
36	1247.5	46.7	2420	9 HSCYCTCT	X58631 Rat mRNA fo
37	1243.5	46.6	2125	10 RPTYKI	U05247 Mus musculu
38	1227.5	46.0	1779	10 MMU05247	AY007162 Homo sapi
39	1217.5	45.6	1682	5 APO52430	AF052430 Xenopus I
40	1217	45.6	1806	5 APO52430	AK113892 Ciona int
41	1057.5	39.6	2205	3 AK113892	AB006558 Ephydalia
42	1004.5	37.6	1733	3 APO67775	AF067775 Hydra vul
43	976.5	36.6	1566	3 ABO96875	AB096875 Caenorhab
44	873.5	32.7	2504	3 ABO96875	
45	838.5	31.4	110000	2 AC073794_2	Continuation (3 of

RESULT 1

ALIGNMENTS

AR053292 AR053292 1942 bp DNA linear PAT 29-SEP-1999
LOCUS AR053292
DEFINITION Sequence 11 from patent US 5834208.
ACCESSION AR053292
VERSION AR053292.1 GI:5978154
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1942)
AUTHORS Sakano,S.
TITLE Tyrosine kinase
JOURNAL Patent: US 5834208-A 11 10-NOV-1998;
FEATURES Location/Qualifiers
source 1..1942
BASE COUNT 365 a 614 c 652 g 311 t
ORIGIN
Alignment Scores:
Pred. No.: 1.93e-158 Length: 1942
Score: 2671.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 gaps: 0
US-09-977-260-2 (1-507) x AR053292 (1-1942)
QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
DB 208 ATGGGGGGGAGAGCTCTCTGGTTCTCGGGGGGCGATTTCAGCGTGTGATTCGCTGAG 267
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
DB 268 GAACCTCCCGGGGTGAGCCCCCGCTTCCTCGAGCCGGGACCCCTCCCGTCAAGCC 327
QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrIlyCysGluHisThr 60
DB 328 AGGATCCCAAGAGGGGCTGGGGCCCGGGGACCAAGTATCACAATCCGAGCACACC 387
QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaThrIleLeuGluAla 80
DB 388 CGGCCCAAGCGAGGAGCTGGGCTTCGCCAAGGGGACGCTGCTACCATCTGGAGGCC 447
QY 81 CysGluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGluLeu 100
DB 448 TGGGAGAACAAGAGCTGGTACCGCGTCAAGCACCAACGAGTGAAGAGGGGCTGCTG 507
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
DB 508 GCAGCTGGGGGCTGGGGGAGCGGAGGCCCTCCGCAAGCCCAAGCTCACCTCATG 567
QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
DB 568 CCTGGTTCACGCGGAGAGATCTGGGCGGAGAGGCTGTCCAGCAGCTGCGCTCCGAG 627
QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysVal 160
DB 628 GATGGGCTGTCTCTGGGGGAGTCCGGCGCCACCCCGGAGTACGTCCTGCTGCTG 687
QY 161 SerPheGlyArgAspValIleHisTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
DB 688 AGCTTTGGCGGCGACATCCATCCATCCGCTGCTCAGCGGCGGAGCCGCTCACAAATC 747
QY 181 AspGlyAlaValAlaPhePheCysAsnLeuMetAspMetValGluHisTrpSerLysAspLys 200
DB 748 GATGAGGCGCGTGTCTTCTGCACTCACTGAGCATGATGAGGATTCAGCAAGAGCAAG 807
QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
DB 808 GGGGCTATCTGCACCAAGCTGTGTGAGACCAAGAGGAGCAAGGAGGAGGAGGAGG 867

QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
DB 868 GAGGAGCTGGCCAGGCGGCGCTGTACTGACCTGCAGCATTTGACATTTGGAGGACAG 927
QY 241 IleGlyGlnGlyGluPheGlyAlaValLeuGlnGlyLysTrpLeuGlyGlnLysValAla 260
DB 928 ATCGGAGGAGGAGATTGGAGCTGTCTCAGGCTGATGACCTGGGGGCAAAAGGTGCC 987
QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluTrpAlaValMet 280
DB 988 GTGAGAGATATCAAGTGTGATGTGACAGCCAGCCCTCTCTGAGCAGAGGCGCTCATG 1047
QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
DB 1048 ACAGAGATGCAACACGAGACCTGCTGCTCCTGGGCTGATCTCCGACACAGGGGCTG 1107
QY 301 TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgTrpArgGly 320
DB 1108 TCAATGTCTATGAGACAGTGTGACCAAGGCAACCTGTGTGAACTTCTGCGGACCGGGGT 1167
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlyMet 340
DB 1168 CGAGCCCTCGTGAACACCGCTCAGCTCTCAGATTTTCTTGACAGTGGCGGAGGATG 1227
QY 341 GluTrpLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
DB 1228 GAGTACCTGGAGAGCAAGACTTGTGACCGGACCTGTGCGCCGCAACATCTGTGTC 1287
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
DB 1288 TCAGAGAGCTGTGGGCAAGGTCAGGACTTTGGCTTGCCCAAGCGGAGCGGAGGGG 1347
QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
DB 1348 CTAGAGTCAAGCCGGCTGCCCGTCAAGTGAAGCGGCGCCAGGCTCTCAAAACAGGGAG 1407
QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrp 420
DB 1408 TTCACCAAGCAAGTGTGTGTGTGAGTTTGGGGTCTGTCTCGGAGAGGCTTCTCATAT 1467
QY 421 GlyArgAlaProTrpProLysMetSerLeuLysGlyValSerGluAlaValGluLysGly 440
DB 1468 GGAAGGCTCTGATCCCTAAATGTCATGAAAGAGTGTGCGAGCCGTGGAGAAAGGG 1527
QY 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerSerCys 460
DB 1528 TACCGCATGAAACCCCGGAGGCTGTGCGAGGCCCGCTGCACTCATGAGACAGCTGC 1587
QY 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
DB 1588 TGGGAGGCAAGCCCGCCCGCCGCAACCTTCCGAAACTGGCGAGAGAGCTGGCCCG 1647
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
DB 1648 GAGCTACGCAAGTCAAGTGTGCGCCAGCTCTGCTCAGGGGAGAGACCGGAGCTCCACC 1707
QY 501 SerProArgSerGlnGluPro 507
DB 1708 TCGCCCCGAAGCCAGAGGCC 1728
RESULT 2
LOCUS E10345 1942 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding human novel tyrosine kinase.
ACCESSION E10345
VERSION E10345.1 GI:22027175
KEYWORDS JP 1995313157-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1942)
AUTHORS Sakano,S.
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

TITLE NEW TYROSINE KINASE
JOURNAL Patent: JP 1995313157-A 1 05-DEC-1995;
COMMENT ASAHII CHEM IND CO LTD
OS Homo sapiens (human)
PN JP 1995313157-A/1
PD 05-DEC-1995
PF 11-AUG-1994 JP 1994189444
PR 25-AUG-1993 JP 93P 210403, 29-MAR-1994 JP 94P 58553 PI
SAKANO SEIJI
PC C12N9/12, A61K38/45, C07K14/47, C12N15/09, (C12N9/12, C12R1.91), PC
(C12N15/09,
C12R1.91);

CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..1942
FT /organism='Homo sapiens'
FT /cell_type='megakaryoblast'
FT /cell_line='UT-7'
FT 5'UTR 1..207
FT CDS 208..1731
FT /product='human novel tyrosine kinase' FT
FT 3'UTR 1732..1942
FT CDS 331..1731
FT /product='human novel tyrosine kinase' FT
FT misc_RNA 349..540
FT /product='SH3 domain'
FT 571..795
FT /product='SH2 domain'
FT misc_RNA 904..1641
FT /product='tyrosine kinase domain'
FEATURES
source 1..1942
/organism='Homo sapiens'
/mol_type='genomic RNA'
/db_xref='taxon:9606'

BASE COUNT 365 a 614 c 652 g 311 t
ORIGIN

Alignment Scores:
Pred. No.: 1.93e-158 Length: 1942
Score: 2671.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-977-260-2 (1-507) x E10345 (1-1942)

OY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
DB ATGGCGGGGCGAGGCTCTCTGTTTCCTGCGGGGATTTCACGGCTGATTCGTGCTAG 267
OY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProProValSerAla 40
DB GAACCTCCCGGGGTAGCCCGCTCTCTCCGAGCTGGACCCCGCTCCGCTCAGCC 327
OY 41 ArgMetProTrpArgArgTrpAlaProGlyThrGlnCysIleTrpLysCysGluHisThr 60
DB AGGATGCCACAGAGGCGCTGGCCCGGCGCACCAAGTATACACCAATGCGACACACC 387
OY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValValThrIleLeuGluAla 80
DB CGGCCCAAGCCAGGGGAGCTGCTTCGCGCAAGGCGACGTCACCATCTCGAGGCC 447
OY 81 CysGluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGluGlyLeuLeu 100
DB TGCAGACAACAGACCTGTCACGCGTCAAGACACACACAGTGTGACAGAGGAGGCTG 507
OY 101 AlaIleGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120

DB 508 GCAGCTGGGGCGCTCGGGAGGGAGGCCCTCTCCGAGAACCCCAACCTACGCTCATG 567
OY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
DB CCGGTTCACAGGGAGATCTCGGGCCAGAGGCTGTCCAGACGCTCAGCTCCCGAG 627
OY 141 AspGlyLeuPheLeuValArgGluSerAlaTrpHisProGlyAspTrpValLeuGlyVal 160
DB GATGGGCTGTTCTGCTGGGAGTCCCGGCGCACCCCGGAGTACGCTGTCGCTG 687
OY 161 SerPheGlyArgAspValIleHisTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
DB AGCTTGGCGCGAGCATCATCAGCGCGTGCACCGCGACCGGCGGCGGCTCATCATC 747
OY 181 AspGluAlaValPhePheCysAsnLeuMetLysPheValGluHisTrpSerLysAspLys 200
DB GATGAGGCGGTCTCTTCGCAACCTCATGAGCATGTGAGACATTACAGCAGAGACAG 807
OY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
DB GCGCTATCTGCACCAAGCTGTGAGACCAAGGGAACACGGAGACCAAGTCCGCGAG 867
OY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
DB GAGGAGCTGGCGAGGGCGGCTGTTACTGACACCTGCAGCATTTGACATTGGAGCACAG 927
OY 241 IleGlyGlnGlyGluPheGlyAlaValLeuGlnGlnGlyLysGlyGlnLysValAla 260
DB ATCGAGAGGAGAGATTGTGAGCTCTCTCGAGGCTGAGTACCTGGGCGCAAAAGGTGCC 987
OY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
DB GTGAGGAATATCAAGTGTGATGTGACACCCAGGCTCTCTCGAGAGACGCGCTCATG 1047
OY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
DB ACGAAGATGCAACAGAGAACCTGTGCTCTCTGGCGCGTGCATCTGACAGAGGGCTG 1107
OY 301 TyrIleValMetGluHisValSerLysGlyLysLeuValAsnPheLeuArgTrpArgGly 320
DB TACATGTGATGAGGACAGCTGAGCAAGGCAACCTGTGATCTTCTCGGACCCGGGGGT 1167
OY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMet 340
DB CGAGCCCTCGTGAACACCGCTCAGCTCTGCAAGTTTCTCTGCAAGTGGCGAGGAGCATG 1227
OY 341 GluTrpLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
DB GAGTAACTGGAGACAGAAAGCTGTGACCGCGACCTGGCGCGCCCGCAACATCTGTGTC 1287
OY 361 SerLysPheLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
DB TCAAGAGACCTGTGGCCAAAGTGTGACGACTTGTGGCTGGCGAAAGCCGAGCGGAAGGG 1347
OY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
DB CTAAGCTCAACCGCGTCCCGCTCAAGTGAAGCGCGCCGAGGCTCTCAAAACAGCGGAAG 1407
OY 401 PheTrpSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrp 420
DB TTCACAGCAAGTGTGATGTCTGAGATTGTGGGCTGCTGCTGGAGAGCTTCTCATAT 1467
OY 421 GlyArgAlaProTrpProLysMetSerLeuLysGlyValSerGluAlaValGluLysGly 440
DB GAGCGGCTCGTACCTTAATAATGTCACTAAAGAGGTGTGAGGCGGTGGAGAAAGGGG 1527
OY 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerSerCys 460
DB TACCGCATGGAACCCCGGAGGCTGTCCAGGCGCCCTGTGACGTCCTCATAGACAGCTGC 1587
OY 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480

Db 1588 TGGGAGGACGACCCCGCCGCGCAACCTCCGCAAGAGAGCTGGCCCGG 1647

QY 481 GtLeuA1ArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
|||||
Db 1648 GAGCTACGACGATGTCAGGTGCGCCAGCCCTCGTCTCAGGGGAGAGAGCCGACGCTCCACG 1707
|||||

QY 501 SerProArgSerGlnGluPro 507
|||||
Db 1708 TCGCCCGGAGAGCCGAGAGCC 1728
|||||

RESULT 3

LOCUS HSHYTK 1968 bp mRNA linear PRI 10-OCT-1994

DEFINITION H.sapiens HYL tyrosine kinase mRNA.

ACCESSION X77278

VERSION X77278.1 GI:471312

KEYWORDS HYLTK gene; nonreceptor protein tyrosine kinase.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Sakano,S., Iwama,A., Inazawa,J., Ariyama,T., Ohno,M. and Suda,T.
TITLE Molecular cloning of a novel non-receptor tyrosine kinase, HYL
JOURNAL (hematopoietic consensus tyrosine-lacking kinase)
MEDLINE Oncogene 9 (4), 1155-1161 (1994)
94181267
PUBMED 8134117
REFERENCE 2 (bases 1 to 1968)
AUTHORS Iwama,A.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-1994) A. Iwama, Dept of Cell Differentiation,
Inst of Mol Embryology & Genetics, Kumamoto University School of
Medicine, 2-2-1 Honjo, Kumamoto 860, JAPAN
Location/Qualifiers
1. 1968
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19p13"
/tissue_type="hematopoietic cell"
/clone_lib="UT-7 cDNA library"
1. 1968
/gene="HMYTK"
1. 1968
/gene="HMYTK"
208. 1731
/gene="HMYTK"
/gene="HMYTK"
/product="HYL tyrosine kinase"
/protein_id="CA54493.1"
/db_xref="GI:557272"
/db_xref="SWISS-PROT:P42679"
/translation="MAGRSLVSMRAFHGCDSSAEELPRVSPFLRAMHPVRSARMT
RRAPDGTCTICRCEHTRPKPGELAPKGDVNTLLACENKSKYRYKHNHSGOGLIA
GALRRERASAPKSLIMPFHGKISGDAVOLODPENGLELVRESANHPGDIYLCV
SEGRDVIHYRVLRDGHITDAVEFCNLMDEHYSKQGAICTKLVPKRKHGKS
AAEELARAGMLNLOHITLGAOIGEGEFAVLQGEYIGQKAVKNTKICVTAQAFDE
TAVMTKMOHEENVRLIGVILHOGILYVMEHVSQNLVNLRTGRALVTAQLOFSL
HYAEGMEYLSKKLVLRDLAARNILYSEDLVAKVSDPGLAKAERGLDSRLPVKMTA
PEAKHGKTSKSDVMSRGVILMEVSYRARYPKSKLSYESEAEVKGRMPREPCP
GPNVHLMSSCWAEPRAPRPFFKRLAEKLRLKLSAGAPASVSGDDADGSTSPSDP"

BASE COUNT 391 a 614 c 652 g 311 t

ORIGIN

Alignment Scores:

Pred. No.: 1,966-158

Score: 2671.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 100.00%

Length: 1968

Matches: 507

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-977-260-2 (1-507) x HSHYTK (1-1968)

QY 1 MetaLag1ArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
|||||

Db 208 ATGGCGGGGCGAGGCTCTGTGGTTCTCTGGCGGGCATTTTCAGGGCTGTGATTCGTGTAG 267
|||||

QY 21 GtLeuF0ARgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
|||||

Db 268 GAACCTCCCCGGGTGAGCCCCCGCTTCCTCCGAGCTCGACCCCCCTCCCTCAGCC 327
|||||

QY 41 ArgMetProThrArgArgTrpAlaProGlyTrpGlnCysIleThrCysGlyHisThr 60
|||||

Db 328 AGATGCCACAGAGGGGCTGGGCGCCCGGACCCAGTATATACCAATGCGAGCACACC 387
|||||

QY 61 ArgProLysProGlyGluLeuAlaPheArgGlyGlyAspValValThrIleLeuGluAla 80
|||||

Db 388 CGGCCCAAGCCAGAGGAGGTGGCTTCCGCAAGAGGCGAGCTGTACATCTGTGAGGCC 447
|||||

QY 81 CysGluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGluLeu 100
|||||

Db 448 TCGGAGACACAGAGCTGTACCGCCGTCACAGCACACACAGTGACAGAGGGGCTGTG 507
|||||

QY 101 AlaAlaGlyAlaLeuArgGluArgGluArgGluAlaLeuSerAlaAspProLysLeuSerMet 120
|||||

Db 508 GCAGCTGGGGCGCTGGGAGCGGAGGCGGCTCTCCGCGAGACCCCAAGCTCAGCTCATG 567
|||||

QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProGlu 140
|||||

Db 568 CCGTGTTCACAGCGAGATCTCGGCGCAGAGAGCTGTCCACAGCTGCAGCTCCGCGAG 627
|||||

QY 141 AspGlyLeuPheLeuValArgGluSerAlaAlaGlnHisProGlyAspTrpValLeuGlyVal 160
|||||

Db 628 GATGGCTGTCTCTGTGGGAGAGTCCGCGCCGACCCCGGAGACACTCTGTGCGTG 687
|||||

QY 161 SerPheGlyArgAspValIleHisTrpArgValIleHisArgAspGlyHisLeuThrIle 180
|||||

Db 688 AGCTTTGGCGCGGACGCTCATCCACTCCGCTGTCTGACCGGAGCGCCACCTCACAATC 747
|||||

QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTrpSerLysAspLys 200
|||||

Db 748 GATGAGCGCGGTGTTCTTGCAACCTCATGAGCATGTGAGACATTCAGCAAGGCAAG 807
|||||

QY 201 GlyAlaIleCysThrTrpLysLeuValArgProLysArgGlyHisGlyThrLysSerAlaGlu 220
|||||

Db 808 GGGCTATGTGCACCAAGTGTGTGAGACCAAGGCAAGCGGAGCAAGTGGCGGAG 867
|||||

QY 221 GtGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
|||||

Db 868 GAGGAGCTGGCCAGGCGGGCTGTGTTACTGMACTCGACATTGATGGAGACAG 927
|||||

QY 241 IleGlyGluGlyGluPheGlyAlaValLeuGlnGlyGlyTrpLeuGlyGlnLysValAla 260
|||||

Db 928 ATCGGAGAGGAGAGATTGTGAGCTGTCTGACAGGGTGATCACTGGGCAAAAGTGCC 987
|||||

QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
|||||

Db 988 GTGAAGAATATCAAGGTATGTGACAGCCAGGCGCTTCTGTGACAGAGCGCCGTCAAG 1047
|||||

QY 281 ThrLysMetGlnHisGluAsnLeuValArgGlyLeuGlyValIleLeuHisGlnGlyLeu 300
|||||

Db 1048 ACGAAGATCAACACAGCAACTGTGGCTCTCTCGGGCTATCTCTGACCAAGGGGCTG 1107
|||||

QY 301 TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
|||||

Db 1108 TACATTGTGATGAGACGTGAGAGCAAGGCAACGTGTGTAACCTTCTGGGAGCCCGGGGT 1167
|||||

QY 321 ArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGluGlyMet 340
|||||

Db 1168 CGAGCCCTGTGTACACCGCTCAGCTCCGAGATTTCCTCAGCTGGCCAGAGGCAAG 1227
|||||

QY 341 GtUtyrLeuGlnSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
|||||

Db 1228 GAGTACTGTGAGAGCAAGAGCTTGTCACCGCGAGACTGTGGCCGCGCAACATCTTGCTG 1287
|||||

QY 361 SerGluAspLeuValAlaIalysValSerAspPheGlyLeuAlaIalysGluArgLysGly 380
 Db 1288 TCACAGGACCTGTGGTGGCCAGAGTCAAGCTTTGGCTGGCCAAAGCGAGCGAGGGG 1347
 QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
 Db 1348 CTACACTCAACCCGGCTGCCCTCAAGTGGAGCGGCCCGGAGCTCTCAACACAGGGAG 1407
 QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerThr 420
 Db 1408 TTCACAGCAGAGTGGATGTCTGGAGTTTGGGGTGGCTGTGGAGAGCTTCTCATAT 1467
 QY 421 GlyArgAlaProLysTrpProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
 Db 1468 GGAGGGGCTCCGTACCCCTAAATGTCACTGAAAGAGGTGTGGAGCGCGTGAAGAAGGGG 1527
 QY 441 TyrArgMetGluProGluGlyCysProGlyProValHisValLeuMetSerSerCys 460
 Db 1528 TACCGCATGGAACCCCGGAGGGCTGTCCAGGCCCGCTGCACGTCTCATAGCAGCTGC 1587
 QY 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
 Db 1588 TGGGAGGACAGACCCCGCGCGGCGCACCTTCGCAAACTGCCGAGAACCTGGCCCGG 1647
 QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyHisAspAlaAspGlySerThr 500
 Db 1648 GAGTACCGCAGTGCAGGTGGCCCGGACCTCCGTCTCAGGGCAGAGCGCGAGGCTCCACC 1707
 QY 501 SerProArgSerGluGluPro 507
 Db 1708 TCGCCCGGACGACGAGAGCC 1728
 RESULT 4
 BDI76710 1997 bp DNA linear PAT 18-MAR-2003
 LOCUS BDI76710
 DEFINITION Method of examining allergic disease.
 ACCESSION BDI76710.1 GI:29122420
 VERSION WO 02075304-A/11.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1997)
 AUTHORS Sugita,Y., Heishi,M., Kagaya,S., Gunji,S. and Tsujimoto,G.
 TITLE Method of examining allergic disease
 JOURNAL Patent: WO 02075304-A 11 26-SEP-2002;
 GENOM RESEARCH INC, JAPAN AS REPRESENTED BY GENERAL DIRECTOR OF NATIONAL CHILDREN'S HOSPITAL, HIDEOTO YONEKURA, YASUHIKO
 YAMAMOTO, SHIGERU SAKURAI, TAKIO MATANABE YUJI SUGITA, MASAYUKI
 HEISHI, SHINJI KAGAYA, SHIGEMICHI GUNJI, GOZO TSUJIMOTO
 OS Homo sapiens (human)
 PN WO 02075304-A/11
 PD 26-SEP-2002
 PR 01-MAR-2002 WO 2002JP001916
 PI 21-MAR-2001 JP 01P 081028
 PI YUJI SUGITA, MASAYUKI HEISHI, SHINJI KAGAYA, SHIGEMICHI GUNJI, PI
 GOZO TSUJIMOTO
 PC GOIN33/15,GOIN33/50,C12N15/09,C12Q1/02,C12Q1/68,A61K45/00, PC
 A61K31/711,
 PC A61K39/395,A61P37/08,A61K48/00,A01K67/027
 CC Method of examining allergic disease
 FH key Location/Qualifiers
 FT CDS Location/Qualifiers
 FEATURES
 source 1..1997
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 372 a 631 c 676 g 318 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2e-158 Length: 1997
 Score: 2671.00 Matches: 507
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-977-260-2 (1-507) x BDI76710 (1-1997)
 QY 1 MetaAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
 Db 266 ATGGCGGGGAGAGCTCTCGTCTCTCCGTGGGGGAGCATTCACGGCTGTGATTCGTGAG 325
 QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProProValSerAla 40
 Db 326 GAACTTCCCGGGGAGACCCCGGCTTCCGAGCCCTGGACCCCGGACCCCTCCGCTTCAGCC 385
 QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
 Db 386 AGGATGCCAACAGAGGCTGGGCGCCCGGACCCCAAGTGTATCACCAAATCGAGACACACC 445
 QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaTrpIleLeuGluAla 80
 Db 446 CGCCCAAGCCAGGGAGAGCTGCTCCGCAAGGGGACGTCGACCATCTCGAGGCC 505
 QY 81 CysGluAsnLysSerTrpTyrArgValLysHisHisThrSerGlyGlnGluGlyLeu 100
 Db 506 TGCAGAAACAAGACTGCTACCGGCTCAAGCACACACACCATGTGACAGAGGAGGCTGTG 565
 QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 Db 566 GCAGCTGGGGGCGCTCGGGAGAGCGGAGGCCCTCTCCGACAGCCCAAGCTCAGCTCATG 625
 QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuProProGlu 140
 Db 626 CCGTGTTCACCGGAGAGATCTCGGGCAGAGGCTGTCCAGCAGCTGCACCTCCCGAG 685
 QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuCysVal 160
 Db 686 GATGGGCTGTTCGGTGGGAGTCCGGGCGCACCCCGGCGACTACGTCTTGCGTG 745
 QY 161 SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle 180
 Db 746 AGCTTGGCCGCGACGTCATCTACCTACCGCTGTGCACCGCGAGCGGCACCTCACAATC 805
 QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLysAspLys 200
 Db 806 GATGAGCGCGTGTCTCTGCAACCTCATGACATGTGTGGAGCATTTACAGCAGACAGAG 865
 QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
 Db 866 GGCGCTATCTCCACCAAGCTGTGAGACCAAGGGAGAAACCGGAGACCAATCGGCGAG 925
 QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
 Db 926 GAGAGCTGGGCGAGGGCGGGCTGTGTTACTGAACCTGCAGCATTTGACATTGGAGCAGAG 985
 QY 241 IleGlyGluGlyLeuPheGlyAlaValLeuGlnGlnGlyLysTrpLeuGlnLysValAla 260
 Db 986 ATCGGAGAGGAGAGATTGGAGCTGTCTCGCAGGCTGTAGTACCTGGGGCAAAAGGTGCC 1045
 QY 261 ValLysAsnIleLysCysAspValTrpThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
 Db 1046 GTGAAGATATCAAGTGTGATGTGACACCCAGGCTTCTCGAGAGAGAGCGGCTGTCTG 1105
 QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyAlaIleLeuHisGlnGlyLeu 300
 Db 1106 AGCAAGATGCAACACAGAACCTGTGTCTCTGGGCGTGTATCTGCACACAGAGGCTG 1165
 QY 301 TyrIleValIleMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
 Db 1166 TACATTGTTCATGTGAGCAGCTGAGCAAGGCAACCTGTGTGAACTTTCTGCGAGCCGGGCT 1225

QY	321	ArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGluGlyMet	340
Db	1226	CGAGCCCTCGGAACACCCGCTACGCTCTGTGCAAGTTTCTCTGCAAGTGGCCGAGGCGCATG	1285
QY	341	GLuTYrLeuGlnSerLysLysLeuValHisArgAspLeuAlaIleArgAsnIleLeuVal	360
Db	1266	GAGTRCCTCGGAGACAGAACCTTGTCACCGCAGCCTGGCCGCCGCAACATCCTGGTCTC	1345
QY	361	SerGlnAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly	380
Db	1346	TCAGAGGACCTGGTGGCCCAAGTCAGGACACTTTGGCCTGGCCCAAGCCGAGCGAAGGGG	1405
QY	381	LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys	400
Db	1406	CTAGACTCAAGCCCGCTCCCGCTCAAGTGAAGCGGGCCGAGGCTCTCAAAACACGGGAG	1465
QY	401	PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyr	420
Db	1466	TTCAACACCAAGTGGAGATCTGTGAGTTTGGGGTGGTCTGTGGAGAGTCTTCTCATAT	1525
QY	421	GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly	440
Db	1526	GGAGCGGCTCCGTACCCCTAAATGTCATGAAAGAGGTGTGCGAGGCGGTGGAAGAGGG	1585
QY	441	TyrTrpMetGluProProGluGlyLysProGlyProValHisValLeuMetSerSerCys	460
Db	1586	TACCGCATGGAACCCCGGAGGGCTGTCAAGCCCCCGGCAAGCTCTCATAGACGCTGC	1645
QY	461	TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg	480
Db	1646	TGGAGGGCAGAGCCCGCCGCGCCGCAACCTTCGCAAACTGGCCGAGAACACTGGCCCGG	1705
QY	481	GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyLysAspAlaAspGlySerThr	500
Db	1706	GAGCTACGACATGGAGAGTGGCCCAAGCCTCCCTCAGAGGGCAGAGCGCAGCGCTCCACC	1765
QY	501	SerProArgSerGlnGluPro 507	
Db	1766	TCGCCCGCAGACCAAGAGGCC 1786	
RESULT 5			
LOCUS	ARI79662	2000 bp	DNA linear PAT 20-APR-2002
DEFINITION	Sequence 1 from patent US 6326469.		
ACCESSION	ARI79662		
VERSION	ARI79662.1	GI:20221217	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 2000)		
TITLE	Ullrich,A., Gishitzky,M. and Sures,I.		
JOURNAL	Megakaryocytic protein tyrosine kinases		
FEATURES	Patent: US 6326469-A 1 04-DEC-2001;		
	Location/Qualifiers		
source	1..2000		
BASE COUNT	375 a 632 c 673 g 320 t		
ORIGIN	/organism="unknown"		
Alignment Scores:			
Pred. No.:	2e-158	Length:	2000
Score:	2671.00	Matches:	507
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-977-260-2 (1-507) x ARI79662 (1-2000)			
QY	1 MetalaglyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20		

Dd	258	ATGGCGGGGCGAGGCTCTCTGGTTTCTTGGCGGGGCAATTTACAGGCTGTGATTTCTGTGAG	317
Qy	21	GIuIeuProAryValSerProAryPheUaAryAlaTrpHisProProValSerAla	40
Dd	318	GAACGTTCCTCCGGGTAGGCCCTCCGTTCTCTCGAGAGCTGGACACCCCTCCGTTCAGGC	377
Qy	41	ArgMetProThrArgArgTTPAlaProGlyThrGlnCysIleThrLysCysGlnHisThr	60
Dd	378	AGAGAGCCACAGAGGCGCTGGGGCCCCGGGGACCCAGTGTATCCAAATATGCAGCAGCACCC	437
Qy	61	ArgProLysProGlyGIuIeuAlaPheArgLysGlyValSerValThrIleLeuGlnAla	80
Dd	438	CGCCCCAGCCAGGGGGAGCTGGCTTTCGGCAAGGGCGAGCGTGTAACCATCTTGGAGGCC	497
Qy	81	CysGIuAsnLysSerTPRTPYArgValLysHisHisThrSerGlyGlnGlnGlyLeuLeu	100
Dd	498	TGGCGAACAAGAGCTGTGTACCGGCTCAAGACACACACAGTGGACAGAGAGGCTGTG	557
Qy	101	AlaAlaGlyAlaLeuArgGluArgGlnAlaLeuSerAlaAspProLysLeuSerLeuMet	120
Dd	558	GCAGCTGGGGCGCTCGGGAGGGGAGGCGCTTCTCGGAGACCCCAAGCTCAGCTCATG	617
Qy	121	ProTPRPhnISGLysIleSerGlyGlnGlnAlaValGlnGlnLeuGlnProProGly	140
Dd	618	CCGTGTGTCCAGGGAGAAATCTCGGGGCCAGAGGCTGTCCAGACATGCAAGCTCCCGAG	677
Qy	141	AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTPYValLeuCysVal	160
Dd	678	GATGGCCCTTTCCTGTGCTGGGGAGTCCGGCCGACCCCGCGAGTACGTCTGTGGCTG	737
Qy	161	SerPheGlyArgAspValIleHisThrArgValLeuHisArgAspGlyHisLeuThrIle	180
Dd	738	AGCTTTGGCCCGACGATCATCCATCAACCGCGTGGCAACCGGACGCGCCACTCCACAAATC	797
Qy	181	AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLysAspLys	200
Dd	798	GATGAGGGCGGTGTCTTCTTGGCAACCTCATGGACATGGAGGAGCATTTACAGCAAGAGCAAG	857
Qy	201	GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu	220
Dd	858	GGCGGTATCTGCACCAAGCTGTGTAGACCAAGCGGAACACGAGGCCCAAGTGGCCGAG	917
Qy	221	GluGluLeuAlaArgAlaGlyTPRLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln	240
Dd	918	GAGGAGCTGGCCAGGCGGGCTGTGTACTGAACCTGCAGCAATTTGACATTGGGAGCACAG	977
Qy	241	IleGlyGlnGlyGlnPheGlyAlaValLeuGlnGlnGlyTyrTyrLeuGlyGlnLysValAla	260
Dd	978	ATCGGAGAGGGAGAGTTTGTGAGCTGTCTCTGAGGGGTGATGACTCGGGGCAAAAAGGTGGCC	1037
Qy	261	ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGlyThrAlaValMet	280
Dd	1038	GTTGAAGATATCAAGTGTGATGTACACGCCAGAGCTTCTGTGACGAGAACGGCTCATG	1097
Qy	281	ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu	300
Dd	1098	ACGAAGATGCAACACGAGAACCTGGTGGCTCTCTGGCGCGATGCTGCACAGAGGGCTG	1157
Qy	301	TyrTlleValMetGlnHisValSerLysGlyLysnLeuValAsnPheLeuArgThrArgGly	320
Dd	1158	TACATTGTTCAGAGGAGCAGTGTAGCAAAAGGGCAACCTGTGTGAATCTTCTCGGAGCCGGGT	1217
Qy	321	ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlnGlyMet	340
Dd	1218	CGAGCGCTCGTGAACACCGCTCAAGCTCTGTGAGTTTCTTGTGCACGTGTGGCGAGGCAATG	1277
Qy	341	GluTyrLeuGlnIuSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal	360
Dd	1278	GAGTACCTGGAGAGCAGAAACCTTGTGTGACCGCGACCTGGGCGCCGCAACATCTGTGTC	1337
Qy	361	SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly	380
Dd	1338	TCAAGAGCACTGTGTGGCCAAAGTCAAGCACTTGTGGCTGGCCAAAGCCGACGCGAAAGGG	1397

FEATURES	SOURCE
<p>FEATURES</p> <p>Location/Qualifiers</p> <p>1. 2180</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="locusid:4145"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="MGC:1708 IMAGE:3350365"</p> <p>/tissue_type="Eye, retinoblastoma"</p> <p>/clone_id="NH_MGC_16"</p> <p>/lab_host="DH10B-R"</p> <p>/note="Vector: pOTB7"</p> <p>400. 1923</p> <p>/codon_start=-1</p> <p>/product="megakaryocyte-associated tyrosine kinase"</p>	<p>CDS</p>

	BASE COUNT	436 a	689 c	724 g	331 t	ORIGIN
						/protein_id="AAH00114.1" /db_xref="GI:12652729" /translation="MAGRSGLVSNAPAFPHGCDSEAFELPSRPFLRAKHPPPVSAAMP RRMAPGCTCCKCENTRPRKPGELARFKDQDVTTILEACKESMTYRKHHNTSGOGLTA GALKREALISADPKLSTMPWPHGKTSGEAVQOOLPDEGLELVESARHPDVLIC SEGRDIYHYRYVLRHGDHLJTIDEAVFPCMLMVAEYYSKDKGATCYKLVAPKKHGTK AEEELIARAGWMLNLNHLTLGAOIGEEFEGVALQGEYLGQVAKVNIKCDVTAQDELDA TAVYTKMOMENLVRLGLYILHQGLYIVNEHYSKGLVNFLEFRGALVNTADLLOFSS HVAEEMYLESKLKHRLDAKRLNLIASDVHAKVSDFLGARERKGLDSSRLPYVMT PEALKHGKFTSKSDVMSRGVLLMEYFVSYGRAPYPMNSLKEVSEANKEGIRMPPEGC GPVHLWSSCWEAEPRARRPPFRKLLEKRLSAGAFVSYSGDADGSTSPRSQP
	Alignment Scores:					
	Pred. No.:	2, 21e-158	Length:	2180		
	Score:	2671.00	Matches:	507		
	Percent Similarity:	100.00%	Conservative:	0		
	Best Local Similarity:	100.00%	Mismatches:	0		
	Query Match:	100.00%	Indels:	0		
	DB:	9	Gaps:	0		
US-09-977-260-2 (1-507) x BC000114 (1-2180)						
OY	1	MetaIaGIaYarGyISerLeuVaISerTrPaRgaIaPhenISgLYCysaSPSeRaIaGlu	20			
Db	400	ATGGCGGGGCGAGGCTCTCTGTTCTCGGCGGCGCATTCACGGCTGTGATCTCTGTAG	459			
OY	21	GLIueuPtoAaYVaISerProaRgPhelueuRgaIaTrPhISProProVaISeRaIa	40			
Db	460	GAACTTCCCGGGGTAGSCCCCGCTCTCCGAGGCTGAGCACCCCTCCGCTTCACGC	519			
OY	41	ArgmetProthArGaRgTrPaIaProGlyThRInGcYsILeThLyCSaGSLuHISThr	60			
Db	520	AGGATGCCAACGAGGCGCTGGGCCCCGGGACACCCAGTATACCAAAATGGAGCACACC	579			
OY	61	ArgProLySProGlyGluLeuAlaPhenArgLySGIaSPaRvaIaThTILeLueGluAla	80			
Db	580	CGCCCCAAGCCAGGGGAGCTGCGCTTCGCGAAGGGCGAGCTGCATCTTGGAGGCC	639			
OY	81	CysGluASnLySserTrPTyRArgVaILySHISThrSerGlyGInGluGluLeuLeu	100			
Db	640	TGCGAGAACAGAGCTGTTACCGGCTCAACGACACACAGTGGAGAGAGGGCTGCTG	699			
OY	101	AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaSPProLyLeuSerLeuMet	120			
Db	700	GCACCTGGGGGCGCTCGGGAGCGGAGGCGCTCTCCGAGACCCCAACCTCAGCTCATG	759			
OY	121	ProTrPheHISGlyLySILeSerGlyGInGluAlaValGInGluLeuGluProProGlu	140			
Db	760	CCGTGGTTCACGGGAGAGATCTCGGGCCAGAGGCTGTCCAGACCTCAGGCTCCGAG	819			
OY	141	ASpGLyLeuPheLeuValArgGluSeRaIaArgHISProGlyaSPTrVaILeucysVal	160			
Db	820	GATGGCGCTGTCTGTGTCGGGAGTCCGCGGCCACCCCGGCGACTACGTCTGTGGCTG	879			
OY	161	SerPheGlyaRgaSPaRvaIILeHISTrArgVaILeuhISaRgaSPeRgLYHISLeuThrILe	180			
Db	880	AGCTTTTGGCCGCGATCATCCATCCGCTGTGACCCCGCGAGGGCGACCTCATCATC	939			
OY	181	ASpGLuAlaValAPhePheCysASnLeuMetASpMetValGluHISTrSerLySasPlys	200			
Db	940	GATAGAGCCGCTGTCTTCTGTCAACCTATGGACATGTGTGGAGCATTCACGACAGACAAG	999			
OY	201	GLyAlaILeCysThrLySLeuValaRgProLySaRgLySHISGlyThLySeraIaGlu	220			
Db	1000	GGCGCTATCTGCGACCAAGCTGTGTAGACCAAAAGCGGAACACAGGAGACCMACTCGGCGAG	1059			
OY	221	GLIuGLyLeuAlaArgAlaGlyTrIPLeuLeuSnLeuGInHISLeuThLeuGlyAlaGIn	240			
Db	1060	GAGGAGCGTGGCGGCGGCTGTGTATCTGAACCTGCGACCATTTGTACATTTGGAGCACAG	1119			
OY	241	ILeGLyGLuGLyIuPheGlyAlaValaLeuGInGlyIuTrLeuGlyGInLySValaIa	260			

```

Db      1120 ATCGAGAGGAGAGAGTTGGAGCTGTCTCTCAGAGGTAGTACCTGGGGCAAAAGGTGGCC 1179
Qy      261 VallysantileysCysaspvalThralaglnalaphelenuaspouthravalamet 280
Db      1180 GTGAAGATATCAAGTGTGATGTACAGCCAGGCTTCTGTGACGAGAGCGCGGTATG 1239
Qy      281 ThrlysmetGlnHISGluAnleuValArgleuGluYValIleleuHISGlnGlyleu 300
Db      1240 ACGAAGATGCAACACGAGAACCTGTCTCTCTCTGGCGGTGATCTTCACACAGGGGCTG 1299
Qy      301 TyrIleValMetGluHISValSerLysGlyAsnleuValAsnphelenuArgThrArgGly 320
Db      1300 TACATTTCTCAGTGAAGCAGTGAAGGCAAGCACTGTGAACTTTCTGGCAGACCGGGGT 1359
Qy      321 ArgAlaIleuValAsnThrAlaGlnleuLeuGlnPheSerLeuHISValAlaGluGlyMet 340
Db      1360 CGAGCCCTCTGGAACACCGCTCAGCTCTCTGCACTTTCTCTGCACTGGCCGAGGCGCATG 1419
Qy      341 GluTyrleuGluSerLysLysleuValHISArgAspLeuAlaAlaArgAsnIleleuVal 360
Db      1420 GAGTACCTGGAGACAGAGAGCTGTGACCGGCACTGGCGCGCCGACATCTCTGTGTC 1479
Qy      361 SerGluAspLeuValAlaLysValSerAspPheGlyleuAlaLysAlaGluArgLysGly 380
Db      1480 TCAGAGAGCTGTGGTGGCCAGAGTCAAGTCTTGGCTGGCCAAAGCCGAGCGAGAGGGG 1539
Qy      381 LeuAspSerSerArgLeuProValLysThrPheAlaProGluAlaLeuLysHISGlyLys 400
Db      1540 CTAGACTCAAGCCGGCTGCTGCTCAAGTGAAGGCGCGCGGCTCTCAACACAGGAGAG 1599
Qy      401 PheThrSerLysSerAspValTrpSerPheGlyValleuLeuTrpGluValPheSerTyr 420
Db      1600 TTCACACAGCAAGTGCAGTGTGTGAGTTTGGGGTGTGCTGTGAGAGTCTTCTCATAT 1659
Qy      421 GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
Db      1660 GGAGGGGCTCTGTAACCTAAATGTCACTGAAGAGTGTGAGGCGCGTGAAGAGAGGG 1719
Qy      441 TyrArgMetGluProProGluGlyCysProGlyProValHISValleuMetSerSerCys 460
Db      1720 TACCGCATGGAACCCCGGAGGGCTGTCCAGGCGCCGCTGCAGCTCATAGAGAGCTGC 1779
Qy      461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
Db      1780 TGGGAGGCAAGAGCCCGCGCGGCGCACCTTCGGCAACTGGCGGAGAGCTGGCCGG 1839
Qy      481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyLysAspAlaAspGlySerThr 500
Db      1840 GAGCTACGACAGTGCAGGTGCCCGCCAGCTCGCTCAGAGGCGAGGCGGACGCTCCACC 1899
Qy      501 SerProArgSerGlnGluPro 507
Db      1900 TCGCCCGGAGAGCCAGAGGCC 1920

RESULT 8
BC003109 LOCUS 2180 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, megakaryocyte-associated tyrosine kinase, clone
MGC:2101 IMAGE:3357648, mRNA, complete cds.
ACCESSION BC003109
VERSION BC003109.1 GI:13111882
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2180)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (13-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

```

```

REMARK
COMMENT
US
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gqabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Matsja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 6 Row: e Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4505108.
location/Qualifiers
1. 2180
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:4145"
/db_xref="taxon:9606"
/clone="MGC:2101 IMAGE:3357648"
/tissue_type="Eye, retinoblastoma"
/clone_id="NIH_MGC_16"
/lab_host="DH10B-R"
/note="vector: pOTB7"
400. 1923
/codon_start=1
/product="megakaryocyte-associated tyrosine kinase"
/protein_id="AAH03109.1"
/db_xref="GI:13111883"
/translacion="MAGRSGLVSRAPFHGCDSEELPRVSPRELRAMHPPVSRMPT
RRAPGTCITKEHTRPKPELAFKRGDVVTTLEACENKSWRVKHNHTSGOGLLA
GALREKALSDADPKSLMPWFHGRISQEAVOOLPEDEGLFVRESARRHGDVLV
SGFQDVHYVLRHDDHLLTIDEAFPCNLMDVHYSKDGALCTKVRPRKRGTS
AEEELRAGLNLNQLTGAOIGEGEFGAVLGEYGVAVKNTICDVTARAFDE
HVAEYTRKHENLVRILSVITHGCIYIMENYSKNLYNPLRTGRALVYTRQLQPSL
HVAEGMEYLESKLVLRDLARNLTVSEDIYAVKVSDEGLAKERKGLDSSRLPVKWA
PEALKHGRFTSKSDVMSFGVLVAFGRAPYPMKMLKVSSEAVENGYMEPEPGCP
GPVHVLMSCEWSEAPRRPFRKLAELRLSAGAPASVSODADGSTRSPSOBP"
BASE COUNT 436 a 689 c 724 g 331 t
ORIGIN

Alignment Scores:
Pred. No.: 2,21e-158 Length: 2180
Score: 2671.00 Matches: 507
Percent Similarity: 100.00% Conserved: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-977-260-2 (1-507) x BC003109 (1-2180).
Qy      1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHISGlyCysAspSerAlaGlu 20
Db      400 ATGGCGGGGGGAGGCTCTGTGTTCTCTGGCGGGCAATTCACGGCTGTGATCTGCTGAG 459
Qy      21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisPropProValSerAla 40
Db      460 GAATCTCCCGGGGAGACCCCGCTCTCGAGCTCGGACACCCCGCTCCCGTCTAGCC 519
Qy      41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHISThr 60
Db      520 AGGATGCCAAGAGGCGCTGGCGCCGCGGACACCAAGTATACCAATGACGAGACACAC 579

```


QY 61 ArgProLysProGlyLysLeuAlaPheArgLysGlyAspValValThrIleLeuGluAla 80
 DB 580 CCCCCCAAGCCAGGGAGCTGGCTCCGCAAGGGGAGCTGTCACATCTCGAGAGGCC 639
 QY 81 CysGluAsnLysSerTrpTyrArgValLysHisHisThrSerGlyGlnGluLeu 100
 DB 640 TCGGAACACAGAGCTGGTACCGCTCAAGCACACACAGTGGACAGAGGGCTGTG 699
 QY 101 AlaAlaGlyAlaLeuArgLysArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 DB 700 GGAGCTGGGGCGCTGGGGAGGGAGGGCCCTCCGGCAGACCCCAAGCTCAGCCCTCAG 759
 QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProGlu 140
 DB 760 CCGTGGTTCACAGGGAGATCTCGGGCCAGAGGCTGTCCAGCAGCTGCAGCTCCCGAG 819
 QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuGlyVal 160
 DB 820 GATGGCTGTCTCTGGTGGGAGTCCGGCCGACCCCGGAGACTCTCTGTCTGTG 879
 QY 161 SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle 180
 DB 880 ACCTTTGGCCGGAGCTCATTCACATCCGCTGCTGCACCGGAGCGCCACTCAGCATTC 939
 QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLysAspLys 200
 DB 940 GATGAGGCGCTGTCTTCTTGTCAACCTCATGACATGTGAGCATTCAGCAAGAGCAAG 999
 QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
 DB 1000 GGCGGTATCTGCACCAAGCTGTGTGAGACCAAGCGGAAACAGGGGACCAAGTCGGCGG 1059
 QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlnAlaGln 240
 DB 1060 GAGGAGCTGGCCAGGGCGGGCTGTTACGAACTGTGCACATTTGACATTTGGGAGACAG 1119
 QY 241 IleGlyGluGlyLeuPheGlyAlaValLeuGlnGlyGlyTyrLeuGlyGlnLysValAla 260
 DB 1120 ATCGGAGAGGAGAGTGTGGAGCTGTCCGAGGGGAGTACTGGGCGCAAAAGGTGGCC 1179
 QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
 DB 1180 GTGAAGATATATCAAGTGTATGTGACAGCCAGGCTTCTGTGACAGAGCGGCTGATG 1239
 QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuGlyValIleLeuHisGlnGlyLeu 300
 DB 1240 ACGAAGATGCACACAGAGAACTGGTGGCTGTCTGGGGGATGCTGTGACCAAGGGGCTG 1299
 QY 301 TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
 DB 1300 TACATTGTCTATGAGACAGCTGAGCAAGGGCAACTGTGTAACTTTCTGGCGGACCCGGGCT 1359
 QY 321 ArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValIleAlaGlnGlyMet 340
 DB 1360 CAGAGCCCTGTAAACACCGCTCAGCTCTCGAGTTTCTCTGCACAGTGGCCGAGGGGCAAG 1419
 QY 341 GluTyrLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
 DB 1420 GAGTACCTGAGAGCAAGAACTGTGTGACCGCGCAGCTGGCGCGCAACATCTCGTGC 1479
 QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
 DB 1480 TCAGAGGAGCTGTGGCCCAAGGTCAAGCGACTTTGGCTGTGGCCAAAGCCAGCGGAGAGGG 1539
 QY 381 LeuAspSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
 DB 1540 CTAGACTCAAGCCGGCTGCCCTCAAGTGGAGCGGCGCGGAGCTCTCAACACGGGAGG 1599
 QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyr 420
 DB 1600 TTCACACAGCAAGTGGATGTCTGAGATTGTGGGGTGTCTCTGGAGAGTCTTTCATAT 1659

QY 421 GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGlnLysGly 440
 DB 1660 GAGCGGGCTCGCTACCTTAAATGTCAAGAAAGAGTGTCCGAGGCGCTGGAGAAAGGG 1719
 QY 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerGlyCys 460
 DB 1720 TACCGATGAAACCCCCCAGAGGCTGTCCAGGCCCCGTCACGCTCTCATGACAGCTGC 1779
 QY 461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
 DB 1780 TGGAGGCGAGACCCCGCGCGGCGCACCTTCCGCAACTGCGGAGAAAGCTGGCGCGG 1839
 QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAlaAspLysThr 500
 DB 1840 GAGCTACGCGAGTGCAGTGTCCCGAGCTCTGTCTGAGGCGAGAGCGCGAGCTCCAGC 1899
 QY 501 SerProArgSerGlnGluPro 507
 DB 1900 TCGCCCCGAAAGCAGAGGCC 1920
 RESULT 9
 AR053291 1521 bp DNA linear PAT 29-SBP-1999
 LOCUS AR053291
 DEFINITION Sequence 10 from patent US 5834208.
 ACCESSION AR053291
 VERSION AR053291.1 GI:5978153
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1521)
 AUTHORS Sakano, S.
 TITLE Tyrosine kinase
 JOURNAL Patent: US 5834208-A 10 10-NOV-1998;
 FEATURES location/Qualifiers
 source 1..1521
 BASE COUNT 295 a 470 c 507 g 249 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5, 32e-158 Length: 1521
 Score: 2662.00 Matches: 506
 Percent Similarity: 99.80% Conservative: 0
 Best Local Similarity: 99.80% Mismatches: 1
 Query Match: 99.66% Indels: 0
 DB: 6 Gaps: 0
 US-09-977-260-2 (1-507) x AR053291 (1-1521)
 QY 1 MetalaglyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
 DB 1 ATGGGGGGGCGAGGCTCTGTGGTTCTTGGCGGAGATTTACAGGCTGTGATCTGCTGAG 60
 QY 21 GluLeuProAlaArgValSerProArgPheLeuAlaGlnAlaTrpHisProProValSerAla 40
 DB 61 GAACCTCCCGGGGTAGGCCCCCGCTTCTCCAGGCTGGCACCCCTCCGCTGCAGCC 120
 QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGlnHisThr 60
 DB 121 AGGATGCCACAGAGGCGCTGGGGCCCGGCGACCCAGCTGATATCCAAATGCGAGCACAC 180
 QY 61 ArgProLysProGlyLysLeuAlaPheArgLysGlyAspValValThrIleLeuGluAla 80
 DB 181 CCGCCCAAGCCAGGGAGAGTGTCTCCGCAAGGGCGAGAGTGTCCACCATCTTGAGAGCC 240
 QY 81 CysGluAsnLysSerTrpTyrArgValLysHisHisThrSerGlyGlnGluGlyLeu 100
 DB 241 TCGGAGAACACAGAGTGTGACCGCTCAAGCACACACAGTGGACAGAGGGGCTGTG 300
 QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 DB 301 GAGCTGGGGCGCTCGCGGAGGCGGAGGCGCTCTCCGCGAGACCCCAAGCTCAGCTCATG 360

QY	121	ProtrpPheHISGLYSYLISerGluGlnLysValGlnGlnLeuGlnProProGlu	140
Db	361	CCGIGGTTCACCGGGAAGATCTCGGGCCAGGAGGCTGTCAGCAGCTGCACCTCCCGAG	420
QY	141	AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuCysVal	160
Db	421	GATGGGGCGTTCGIGTGGGGAGTCCGGCCGACCCGGCGACTACGTCTGTGGCTG	480
QY	161	SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle	180
Db	481	AGCTTTGGCCCGCAGCGTCATCCACTACCGCGCTGGACCCCGCAGCGCCACTCAATC	540
QY	181	AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLysAspLys	200
Db	541	GATAGAGCCCGTGTCTTCTGCAACCTCATGGACATGGTGGAGCATTACAGCAGACAG	600
QY	201	GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu	220
Db	601	GGCCCTATTCTGCACCAAGCTGGTAGACCAAAACGGGAAACACGGGACCAATCGGCCAG	660
QY	221	GluGlnLeuValArgAlaGlyTyrProLeuLeuGlnGlnHisIleuThrLeuGlyAlaGln	240
Db	661	GAGAGGCTGGCGAGGCGGGGCTGGTTACTGAACTCGCAGCATTTGTACATTGGAGCACAG	720
QY	241	IleGlyGlnGlyGlnPheGlyAlaValLeuGlnGlnGlyLysLysGlyGlnLysValAla	260
Db	721	ATCGAGAGGGAGAGTTGGAGCTGTCTCGCAGGGGTGATGTCGTGGGCAAAAGGTGGCC	780
QY	261	ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet	280
Db	781	GTTGAAGATATCAAGTGTGATGTGACACCCAGGCGCTTCGAGCAGAACCGCCGTCTATG	840
QY	281	ThrLysMetGlnHisGluAsnLeuValArgLeuGlnGlyValIleLeuHisGlnGlyLeu	300
Db	841	ACGAAGATGCACACGAGAACCTGGTGGCTTCTCGGCGTGAATCTGCACAGGGGCTG	900
QY	301	TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly	320
Db	901	TACATTGTCAAGGAGACGCTGAGCAAGGGCAACCTGGTACTTCTCGGACCGCGGGGT	960
QY	321	ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlnGlyMet	340
Db	961	CGAGCCCTCGGAGACACCGCTCACCTCTCTGAGTTTCTTGACAGCTGGCCGAGGGCATG	1020
QY	341	GlyLysLeuGlnSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal	360
Db	1021	GAGTACCTCGAGACGACAAAGACTTGTGACCGCGACCTGGCCCGCCGACATCTGTGTC	1080
QY	361	SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly	380
Db	1081	TCAAGAGGACCTGGGGCCAAAGTCAAGCACATTGGCCCGGCCAAAGCCGACGGAAGGGG	1140
QY	381	LeuAspSerSerArgLeuProValLysThrPThrAlaProGluAlaLeuLysHisGlyLys	400
Db	1141	CTAAGACTCAACCGCGGTCCCGCTCAAGTGAAGGAGCGGCCGAGGCTCTAAACACGGAGAG	1200
QY	401	PheThrSerLysSerAspValTyrSerPheGlyValLeuLeuTyrGluValPheSerTyr	420
Db	1201	TTTCAACAGCAAGTGGATGCTCGAGATTTTGGGGTGGCTCTCGGAGAGCTTCTCATAT	1260
QY	421	GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGlnLysGly	440
Db	1261	GGACGGGCTCCGTACCTTAATAATGTCACTGAAAGAGGTGTGGAGGCCGTGGAGAGGGG	1320
QY	441	TyrArgMetGluProProGlnGlyCysProGlyProValHisValLeuMetSerSerCys	460
Db	1321	TACCGCATGAAACCCCGCAGGGGCTGTCCAAGCCCGCGCAAGTCTCATGAGCAGTGC	1380
QY	461	TyrGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg	480
Db	1381	TGGAGAGGAGAGCCCGCCCGCGCCACACTTCGCAAACTGGCCGGAAGACTGCCCCGG	1440

QY	481	GlueuhtgyserrralagialaProalaserlaserlglycuhnapalaaaglySerThr	500
Db	1441	GAGCTACGACCATGGAGGGCCCGACCTCTGCTTCAGGCGAGACGCCGAGGCTCCACC	1500
QY	501	SerProAargserGingluPro 507	
Db	1501	TCGCCCCGACCGAGAGGCC 1521	
RESULT 10			
LOCUS	S71669	1744 bp	mRNA linear PRI 24-OCT-1994
DEFINITION	1sk-leukocyte carboxyl-terminal src kinase related gene [human,		
ACCESSION	S71669		
VERSION	S71669.1	GI:559593	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1744)		
TITLE	McVicar,D.W., Lai,B.K., Lloyd,A., Kawamura,M., Chen,Y.Q., Zhang,X., Staples,J.E., Ortaldo,J.R. and O'Shea,J.J.		
JOURNAL	Molecular cloning of 1sk, a carboxyl-terminal src kinase (csk)		
MEDLINE	related gene, expressed in leukocytes		
PUBMED	Oncogene 9 (7), 2037-2044 (1994)		
REMARK	GenBank staff at the National Library of Medicine created this entry [NCBI gidsb 151068] from the original journal article.		
FEATURES	This sequence comes from Fig. 1.		
Source	Location/Qualifiers		
gene	1..1744		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	1..1744		
	/gene="1sk"		
	/note="leukocyte carboxyl-terminal src kinase related		
	gene"		
	125..1522		
	/gene="1sk"		
	/note="leukocyte carboxyl-terminal src kinase related		
	gene. This sequence comes from Fig. 1"		
CDS	/protein_id="AAB30995.1"		
	/db_xref="GI:559594"		
	/translation="MPTRRMAPGQCITKCEHTPRPKGELAFRKGDVVTTLACENKSS		
	MYRKHNHSGOEGELLAGALREPLASDPRLSLMPFHGKISGOEAVOOLQDPEDLL		
	FLVSGSGHGDVYLCTGFGDVIHYVLLHDGILLTIDENAFPCNLMDMENVSKDKG		
	ALICRLVLRPKRKHGTSASBEELAAAGMLINQHTLTAAQISGEFGAVLDGEYIGQKC		
	VKNLCVDVTAQAFLEDTAVMTKQHEHNVRLIGVILDQGLYVHEHVKGLVNFLEK		
	RGRLNLTVAQAFLEDTAVMTKQHEHNVRLIGVILDQGLYVHEHVKGLVNFLEK		
	ERKIDISRLRPVKTATADGALKHGFTSKSDWSFGVILMEVFSYGAPYKMSLKEVSS		
	EAVKRYRMPEEGCGPGRVHVMSSCWEAEAPARRPPRLAEKLARLSAGAPASVSS		
	GQDADGSPRSVSOEP".		
BASE COUNT	354 a 534 c 582 g 274 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.62e-154	Length:	1744
Score:	2607.50	Matches:	497
Percent Similarity:	98.62%	Conservative:	6
Best Local Similarity:	98.22%	Mismatches:	1
Query Match:	97.62%	Indels:	1
DB:	9	Gaps:	1
QY	2	AlaaglyAargglySerleuValSerTPPArgAlaPheHISGLYCYSAAPSerAlaGluGlu 21	
Db	5	GCGGGCGCAGGACTCTGCGTTCCCTGGCGGACATTTACGCGCTGTGATTCGTGCAGGAA 64	

QY 22 LeuProArgValSerProArgPheLeuArgAlaTrpHisProProProValSerAlaArg 41
 DB 65 CTTCCTCCGGGTGAGCCCCCTTCCTCCGAGCCGGGACCCCTCCGCTCCAGCCAGG 124
 QY 42 MeProThrArgArgTrpAlaProGlyThrGlnCysIleThrIlyscysGluHisThrArg 61
 DB 125 ATGCCAAGCAGGCGCTGGCCCGGAGCACCAGTATCAACAAATCGAGCACACCAGC 184
 QY 62 ProIlySerProGlyGluLeuAlaPheArgIlysglyAspValAlaThrIleLeuGluAlaCys 81
 DB 185 CCNAAGCAGGAGGAGTGGCCCTCCGCAAGGCGAGCTGGTACACATCTCGAGGCGCTGC 244
 QY 82 GluAsnIlySerTrpTrpArgValIlyshHisIshSerGlyGlnGlyLeuLeuAla 101
 DB 245 GAGAAACAAGAGCTGGTACCCGCTCAACACACACACCTGAGACGGGGGCTGCTGCA 304
 QY 102 AlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProIlySerLeuSerLeuMetPro 121
 DB 305 GCTGGGCGCTGGCGGAGCGGAGGCGCTCTCCGCAAGCCCAAGCTCAGCTCATGCGC 364
 QY 122 TrpPheHisGlyIlyshIlySerGlyGlnGluAlaValGlnGlnLeuGlnProProGluAsp 141
 DB 365 TGGTTCACGAGAAATCTGGGCGACAGAGCTGTCCAGCAGCTGCAGCTCCCGAGAT 424
 QY 142 GlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysValSer 161
 DB 425 GGGCTGTCTCTGTCGAGGAGTCCGGCGGCGCCGACCTCAACCTGCTGTGTCAGC 484
 QY 162 PheGlyArgAspValIleHisIlyArgValLeuHisArgAspGlyHisLeuThrIleAsp 181
 DB 485 TTGGGCGGCGACGTCATACCTACCGGCTCTGACCGGCGACCGCCCTCACAATGAT 544
 QY 182 GluAlaValAlaPhePheCysAsnLeuMetAspMetValGlnIshIlySerIlyAspIlyGly 201
 DB 545 GAGGCGCTGTCTCTCAACCTCATGAGCATGTTGAGCATTAACACAGACAGCAAGGCG 604
 QY 202 AlaIleCysThrIlySerValAlaArgProIlyArgIlyshIshIlyThrIlySerAlaGlu 221
 DB 605 GCTATTCGACACCAAGCTGTGAGACCAACGCGAAGACGCGAAGCTGCGCGAGAG 664
 QY 222 GluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisIshLeuThrIlyGlyAlaGlnIle 241
 DB 665 GAGCTGGCAGGGGCGGCTGTACTGAACTCGACGATTTGACATTTGGAGCAGCATC 724
 QY 242 GlyGlnGlyIlyPheGlyAlaValLeuGlnGlyIlyTrpLeuGlyIlyValAlaVal 261
 DB 725 GGAAGAGGAGAGTGTGGAGCTGTCTGCAAGGTGAGTACCTGGGGCAAAAG---TGGCTG 781
 QY 262 LysAsnIleLysCysAspValIleThrAlaGlnAlaPheLeuAspGluTrpAlaValMetThr 281
 DB 782 AAGAAATATCAAGTGTGATGTGACAGCCCAAGGCTTCCTGAGCAGAGCGCGCTCATGACG 841
 QY 282 LysMetGlnHisGlnAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeuTyr 301
 DB 842 AAGATGCAACACAGAACTGTGCTGTCTCTGGGCTGTATCTCTGACACAGGGGCTGTAC 901
 QY 302 IleValMetGlnHisValSerIlyGlyAsnLeuValAsnPheLeuArgTrpArgIlyArg 321
 DB 902 ATGTGATGAGACGACGAGCAGCAAGGCGCAACCTGTGAACTTCTGCGGAGCCCGGGTGA 961
 QY 322 AlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlnGlyMetGlu 341
 DB 962 GCGCTGTGAACACCGCTGAGCTGTGCACTTTTCTGTGCACTGGCGGGGCGCATAGAG 1021
 QY 342 TyrLeuGlnSerIlySerIlyLeuValHisArgAspLeuAlaAlaArgAsnIleLeuValSer 361
 DB 1022 TACTTGAGAGCAAGAACTTTGTGCACCGGAGCTTACCGCCCGCAACATCTGCTGTCA 1081
 QY 362 GluAspLeuValAlaIlyValSerAspPheGlyLeuAlaIlyValAlaGlnArgIlyGlyLeu 381
 DB 1082 GAGAGAGTGTGGCCCAAGGTCAGGACTTGTGGCTGGCCAAAGCCGAGGAGGGGCTA 1141
 QY 382 AspSerSerArgLeuProValIlyTrpThrAlaProGluAlaLeuIlyshIlyGlyPhe 401

DB 1142 GACTCAAGCCGCGTGCCTCAAGTGACGCGCCCGCAGGCTCTCAAAACAGGGAAGTTC 1201
 QY 402 ThrSerIlySerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrpGly 421
 DB 1202 ACCAGCAAGTCGAGTCTGTGAGTTTGGGCTCTGCTGTGGAGGCTTTCTATATGA 1261
 QY 422 ArgAlaProTrpProIlyMetSerLeuIlysglyValSerGlnAlaValGlnIlyGlyTyr 441
 DB 1262 CGGGCTCCGTAACCTAAATGTCACTGAAGAGGTGTGGAGGCGGTGGAGAGCGGTAC 1321
 QY 442 ArgMetGluProProGlnGlyCysProGlyProValHisValLeuMetSerSerTrp 461
 DB 1322 CGCATGGAACCCCGGAGGCTGTCCAGGCGCGCTGACGCTCTCATAGACCACTGCTG 1381
 QY 462 GluAlaGluProAlaArgProProPheArgIlySerAlaGluIlySerLeuAlaArgGlu 481
 DB 1382 GAGGCAAGCGCCCGCGCGGACACCTTCGCAAACTGGCCGAAAGCTTGCGCGGAG 1441
 QY 482 LeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerTrpSer 501
 DB 1442 CTACGCAAGTCAGAGTGGCCCGACCTCGCTCAGGGGAGAGCCGACGCTCCACTCG 1501
 QY 502 ProArgSerGlnGluPro 507
 DB 1502 CCGCGAAGCCAGAGGCC 1519
 RESULT 11
 LOCUS AR084686 1987 bp DNA linear PAT 01-SEP-2000
 DEFINITION Sequence 1 from patent US 5981201.
 ACCESSION AR084686
 VERSION AR084686.1 GI:10011456
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1987)
 AUTHORS Avraham, H. and Groopman, J.E.
 TITLE Methods of detection and treatment of breast cancer
 JOURNAL Patent: US 5981201-A 1 09-NOV-1999;
 FEATURES
 source location/Qualifiers
 1..1987
 BASE COUNT 369 a 628 c 672 g 318 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,89e-153 Length: 1987
 Score: 2591.50 Matches: 502
 Percent Similarity: 99.21% Conservative: 1
 Best Local Similarity: 99.01% Mismatches: 3
 Query Match: 97.02% Indels: 3
 DB: 6 Gaps: 1
 US-09-977-260-2 (1-507) x AR084686 (1-1987)
 QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
 DB 263 ATGGCGGGGAGAGGCTCTGTGTTCTGTGGCGGCAATTTCACGGCTGTATTTCTGTGAG 322
 QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
 DB 323 GAACCTCCCGGCTGAGCCCGCTTCCTCCGAGCTGGGACCCCGCTCCGCTCAGCC 382
 QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrIlyscysGluHisThr 60
 DB 383 AGGATGCCAAACGAGGCGTGGGCGCGGACCACTGTATCAACCAAAATGCGAGCACACC 442
 QY 61 ArgProIlyProGlyGluLeuAlaPheArgIlysglyAspValAlaThrIleLeuGluAla 80
 DB 443 CGCCCAAGCCAGGAGGAGCTGGCTTCGCAAGGCGACAGTGTCTCACTCTGAGGCC 502

QY 81 CysGluAsnLysSerTrpTyrArgValLysHisHisThrSerGlyGlnGluLeu 100
 |||||
 Db 503 TGGAGAAACAAGAGCTGTACCGGCTGACAGACACACAGTGGACAGAGGGCTCTG 562
 QY 101 AlAlaGlyAlaLeuArgGluArgGluAlaLeuSerLysAspProLysLeuSerLeuMet 120
 |||||
 Db 563 GCAGCTGGGGCGGCGGAGCGGGGAGCCCTCCAGACCCCAAGCTCAGCTCATG 622
 QY 121 ProTrpPheHisGlyLysHisSerGlyGlnGluAlaGlnGlnLeuGlnProProGlu 140
 |||||
 Db 623 CCGTGGTCCACGGGAGATCTGGGCGAGAGGCTGTCCAGCAGCTGAGCCCTCCGAG 682
 QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuGlyVal 160
 |||||
 Db 683 GATGGGCTGTCTCGTGGGAGTCCGGCGGACCCCGGAGCTACGCTCTGTGGTG 742
 QY 161 SerPheGlyArgAspValHisHisTyrArgValLeuHisArgAspGlyHisLeuThrTle 180
 |||||
 Db 743 AGCTTTGGCGCGCGCTCATCCATACCGGCTGCTGACCGCGACCGCCACCTCACATC 802
 QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTyrSerLysAspLys 200
 |||||
 Db 803 GATGAGCCCGTGTCTCTGCAACCTCATGAGCTGTGGAGCTTACACGACAGAGAG 862
 QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
 |||||
 Db 863 GGGCGCTATCGACCAAGCTGGTGAACCAAGGGAGAACAGCGGACCAAGTCCGCGAG 922
 QY 221 GluGluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
 |||||
 Db 923 GAGAGCTGGCCAGGGGCGGCTGTACTACCACTGACGATTTGACATTGGGAGCAG 982
 QY 241 IleGlyLysGluGluPheGlyAlaValLeuGlnGlyLysLysGluLysValAla 260
 |||||
 Db 983 ATGGAGAGGAGAGAGTGGAGCTGTCTGAGGCTGATGACCTGGGAGCAAAAGGTGCC 1042
 QY 261 ValLysAsnLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
 |||||
 Db 1043 GTGAGAAATATCAAGTGTGATGTGACAGCCAGGCTTCCGAGCGAGAGCGGCTCATG 1102
 QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
 |||||
 Db 1103 ACGAAGATGCAACGAGAACCTGTGCTGCTGCGGCTGATCCGACACAGGGCTG 1162
 QY 301 TyrIleValMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
 |||||
 Db 1163 TACATTTCATGAGGACGCTGAGCAGAGGCAACCTGTGTGACTTTCGCGAGACCCGGGCT 1222
 QY 321 ArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGluGlyMet 340
 |||||
 Db 1223 CGAGCCCTCGTGAACACCGCTCAGCTCGTCAAGTTTCTCGACGCGGCGAGGGCATG 1282
 QY 341 GluTyrLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
 |||||
 Db 1283 GAGTACCTGAGAGCAAGAGCTGTGTGACCGGACCTGGCGCCCAACATCTCTGTC 1342
 QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
 |||||
 Db 1343 TCGAGAGACCTGTGGCCAAAGCTCAGGACCTTGGCTGGCCAAAGCGGAGGAGGG 1402
 QY 381 LeuAspSerSerArgLeuProValLysTyrThrAlaProGluAlaLeuLysHisGlyLys 400
 |||||
 Db 1403 CTGAGACTCAAGCGGCGTGGCCGTCAGTGGAGCGCGCGAGGCTCTCAAAACGCGG 1459
 QY 401 PheThrSerLysSerAspValTyrPheSerPheGlyValLeuLeuTyrGluValPheSerTyr 420
 |||||
 Db 1460 TTACACGCAAGTCGATGCTGAGATTGGGCTGTGCTGGGAGGCTCTCTCATAT 1519
 QY 421 GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
 |||||
 Db 1520 GCGAGGCGCTCGTACCTAAATATGTCATGAAGAGGTGTGCGAGGCGCTGGAGAACGGG 1579
 QY 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerSerCys 460

Db 1580 TACCGATGGAACCCCGGAGGGCTGTCCAGGCCCCGCTGACACGCTCATAGAGCACTGC 1639
 QY 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
 |||||
 Db 1640 TGGGAGGCAAGAGCGGCC-CCCGCGCCACCTTCGCAAACTGGCGGAGAGCTGGCCCGG 1698
 QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
 |||||
 Db 1699 GAGCTTACGACAGTGCAGGTGTGCCCGGCTCGTCTCAGGAGGAGAGCGGAGGTTCC-ACC 1757
 QY 501 SerProArgSerGlnGluPro 507
 |||||
 Db 1758 TCGCCCGGAGCCAGAGGCC 1778
 RESULT 12
 AR204704 1987 bp DNA linear PAT 20-JUN-2002
 LOCUS AR204704
 DEFINITION Sequence 1 from patent us 6368796.
 ACCESSION AR204704
 VERSION AR204704.1 GI:21502099
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1987)
 AUTHORS Avraham, H. and Groopman, J. E.
 TITLE Methods of detection and treatment of breast cancer
 JOURNAL Patent: US 6368796-A 1 09-Apr-2002;
 FEATURES
 source location/Qualifiers
 1..1987
 /organism="unknown"
 BASE COUNT 369 a 628 c 672 g 318 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.89e-153 Length: 1987
 Score: 2591.50 Matches: 502
 Percent Similarity: 99.21% Conservative: 1
 Best Local Similarity: 99.01% Mismatches: 3
 Query Match: 97.02% Indels: 3
 DB: 6 Gaps: 1
 US-09-977-260-2 (1-507) x AR204704 (1-1987)
 QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
 |||||
 Db 263 ATGGCGGGGAGAGCTCTGTGTTCTCTGGCGGCGCATTTCAAGCTGTGATTCGTGAG 322
 QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisPProProValSerAla 40
 |||||
 Db 323 GAACCTCCCGGGGTGAGCCCGGCTCTCTCGAGCTGGACACCCCTCCGCTCATGCC 382
 QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
 |||||
 Db 383 AGGATGCCAAGAGAGCGCTGAGCCCGGCGGACCAAGTATACCAAAATGAGACACACC 442
 QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaThrIleLeuGluAla 80
 |||||
 Db 443 CGCCCAAGCAGGAGGAGCTGTGCTCGGAAAGGCGACGTGTGCACATCTGGAGGCC 502
 QY 81 CysGluAsnLysSerTrpTyrArgValLysHisHisThrSerGlyGlnGluLeu 100
 |||||
 Db 503 TGGGAGAACAAAGCTGTACCGCGCTGACACACACACAGTGGACAGAGGGCTCTG 562
 QY 101 AlAlaGlyAlaLeuArgGluArgGluAlaLeuSerLysAspProLysLeuSerLeuMet 120
 |||||
 Db 563 GCAGCTGGGCGCTGCGGAGCGGAGGCGCTCTCCGACAGACCCCAAGCTCAGCTCATG 622
 QY 121 ProTrpPheHisGlyLysHisSerGlyGlnGluAlaGlnGlnLeuGlnProProGlu 140
 |||||
 Db 623 CCGTGGTCCACGGGAGATCTCGGCGAGAGGCTGTCCAGCAGCTGACGCTCCCGAG 682

```
OY 141 AapGlyLeuPheLeuValAArgLysSerAlaArgHisProGlyAspTyrValLeuGlyVal 160
    |||
DB 683 GATGGCTGTTCCTGGTGGGAGTCCGGCCGACCCCGGAGACTACCTCTGTCCTG 742
OY 161 SerPheGlyArgAspValIleHisTyrArgValIleuHisArgAspGlyHisLeuThrIle 180
    |||
DB 743 ACCTTTGGCCGGGAGCTCATCTCACTACCCGCTGCTGACCGGAGCGGACCTCACAATC 802
OY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLysAspLys 200
    |||
DB 803 GATGAGGCGCGTGTCTTCTGCAACCTCATGAGCATGTGAGCATATACAGCAAGGCAAG 862
OY 201 GlyAlaIleCysThrLysLeuValAArgProLysArgLysHisGlyThrLysSerAlaGlu 220
    |||
DB 863 GGGCTATCTGCACCAAGCTGTGAGACCAAGGCGGCAACAGCGGACCAAGTCCGGCCGAG 922
OY 221 GluGluLeuAlaArgAlaGlyTTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
    |||
DB 923 GAGGAGCTGGCCAGGGCGGGCTGCTTACTGAACTGACACATTTGACATTTGGGAGCAG 982
OY 241 IleGlyGluGlyLeuPheGlyAlaValLeuGlnGlyLysLysValAla 260
    |||
DB 983 ATCGGAGAGGAGAGATTGTGAGCTGTCTGCAAGGGTGAAGTACTGGGCGCAAAAGTGGCC 1042
OY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
    |||
DB 1043 GTGAAGAATATCAAGTGTATGTGACAGCCAGGCGCTTCTGACAGCAGCGCCGCTCAG 1102
OY 281 ThrLysMetGlnHisGluAsnLeuValAArgLeuGlyValIleLeuHisGlnGlyLeu 300
    |||
DB 1103 AGGAGATGCACACAGAGAACTGTGTGCTCTCTGGGCTGATCTTCGACACAGGGGCTG 1162
OY 301 TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
    |||
DB 1163 TACATTTGTATGAGAGACGTGAGCAAGGCAACTGTGTAATTTCTGGGAGCCCGGGCT 1222
OY 321 ArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGluGlyMet 340
    |||
DB 1223 CGAGCCCTCTGTAACACCGCTGAGCTCTGCAAGTTTCTCTGACCTGGCCAGGGCAG 1282
OY 341 GluTyrLeuGlySerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
    |||
DB 1283 GAGTACCTGGAGACAGAAAGCTTGTGCACCCGCACTGGCCCGCAACATCTCTGCTC 1342
OY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
    |||
DB 1343 TCAGAGGACCTGTGGCCCAAGGTCAGCGACTTGGGCTGGCCAAAGCCGAGCGAAGGG 1402
OY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
    |||
DB 1403 CTAGACTCAAGCGGCTGCCCTCAAGTGGACGGGCGCCGAGGCTCTCAAAACACGGG 1459
OY 401 PheThrSerLysSerAspValTrpSerPheGlyValIleLeuThrGluValPheSerTyr 420
    |||
DB 1460 TTCACACAGCAATCGGATGCTGGAGTTTGGGGTGCTCTCTGGGAGGTTCTTTCATAT 1519
OY 421 GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
    |||
DB 1520 GGACGGGCTCCTACCTCAATATGTCAGAAAGAGTGTGAGAGCGCTGAGAGAGGGG 1579
OY 441 TyrArgMetGluProProGluGlyLysProGlyProValHisValLeuMetSerCys 460
    |||
DB 1580 TACCCCATGGAACCCCGGAGGGCTGTCCAGGCGCCGCTGCAAGTCTCTCATGAGCTGC 1639
OY 461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
    |||
DB 1640 TGGGAGGCGAGACCCGC-CGCGCGGCACCTTCCGCAAACTGGCCGAGAGAGTGGCCGG 1698
OY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGluAspAlaAspLysSerThr 500
    |||
DB 1699 GAGCTAACGACGACGAGGTGCCCAAGCTCTGCTCAGGCGAGAGCGCCGACGCTCC-ACC 1757
OY 501 SerProArgSerGlnGluPro 507
```

```
DB 1758 TCGCCCGGAGCCAGAGGCC 1778
RESULT 13
HUMMATK
LOCUS Human tyrosine kinase (MATK) mRNA, complete cds.
DEFINITION L18974
ACCESSION L18974
VERSION L18974.1 GI:455449
KEYWORDS cytoplasmic protein; tyrosine kinase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1987)
AUTHORS Bennett,B.D., Cowley,S., Jiang,S., London,R., Deng,B., Grabarek,J.,
Groopman,J.E., Goeddel,D.V. and Avraham,H.
TITLE Identification and characterization of a novel tyrosine kinase from
megakaryocytes
JOURNAL J. Biol. Chem. 269 (2), 1068-1074 (1994)
MEDLINE 94117408
PUBMED 8288563
COMMENT Original source text: Homo sapiens cDNA to mRNA.
FEATURES
source
    location:source text: Homo sapiens cDNA to mRNA.
    1..1987
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /cell_line="CMK11-5"
    /cell_type="megakaryocyte"
    /dev_stage="adult"
    /tissue="lib-lambda gt11"
    1..1987
    /gene="MATK"
    <1..63
    /gene="MATK"
    263..1946
    /gene="MATK"
    /codon_start=1
    /product="tyrosine kinase"
    /protein_id="AA16703.1"
    /db_xref="GI:459550"
    /translation="MAGRSGLYSNPAFHCDSAEELPRVSPFLAAMPBPYSAMPT
    RRMARGTQCTICENTRPKPGELARFGRDQVITLDECEKSWYRVKHNHSGEGLLAA
    GALRGEALADPKLSLMPWFKLSGDEAVOQLOPPEDEGLFLVESARHPQDYLCV
    SEGRDVIHYRLVRDGHLLTIDAVFECNLMDEHYSKDGAICTRLVAPKPHGTS
    ABEELARAGMLNLQHLTGAOIGEGEAGVLOGEYLGOKVAVKNIKCDVQDAFLDE
    TVAVTKMOHENTVRLGVITLHGLITYMEHYSKNLVNFRTGRSALVNTQLDPSL
    HVAEGMETIESKLYLRDILANRNLITYSEDIYAKYSDFLAKAEKGLDSSRLPYKMTA
    PEALKHGFTSKSDVNSFVLLMEVFSYGAPYPKNSLKEVSEAVEKGRIMEPEGCPG
    PVHVLMSGWEAEPPAGHPSANWPMPSWGSYAVOVPPPSOGRTPTVHLAPKPAITP
    FCGPWPORTEREESAAMGH"
    568
    polyA_site
    BASE COUNT 369 a 628 c 672 g 318 t
    ORIGIN
Alignment Scores:
Pred. No.: 1,89e-153 Length: 1987
Score: 2591.50 Matches: 502
Percent Similarity: 99.218 Conservative: 1
Best Local Similarity: 99.018 Mismatches: 3
Query Match: 97.02% Indels: 3
DB: 9 Gaps: 1
US-09-977-260-2 (1-507) x HUMMATK (1-1987)
OY 1 MetaLaglyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
    |||
DB 263 ATCGCGGGGCGAGGCTCTCTGCTTCTGCGCGGCAATTCACAGCCTGATTCGCTGAG 322
OY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
    |||
```

Db 323 GAACCTCCCGGGTGAGCCCCCGCTTCCTCCGAGCCCTGGACACCCCTCCCTCCCTCAGCC 382
 QY ArgMetProThrArgArgTrrPalaProGlyThrGlnCysIleThrLysCysGlnHisThr 60
 Db 383 AGATGTCGAAG 442
 QY ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaThrIleLeuGlnAla 80
 Db 443 CGCCCCAAGCCAG 502
 QY CysGlnAsnLysSerTrrPrrArgValLysHisHisThrSerGlyGlnGlyLeuLeu 100
 Db 503 TGGCAGACACAG 562
 QY AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 Db 563 GCAGCTGGGGCGCTGGCGGGAG 622
 QY ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProGlu 140
 Db 623 CCGGTGTTCCAGAGAGAGATCTCGGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 682
 QY AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrrValLeuCysVal 160
 Db 683 GATGGGCTGTTCTGCTGGAG 742
 QY SerPheGlyArgAspValIleHisTrrArgValLeuHisArgAspGlyHisLeuThrIle 180
 Db 743 AGCTTTGGCGCGCAGATCCTACCTACCGGCTGCTGACAGCGGCGCCACCTCAGCAATC 802
 QY AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTrrSerLysAspLys 200
 Db 803 GATGAGGCGCGTCTCTGCAACCTCATGAGATGGTGAGCATATACAGCAAGAGACAG 862
 QY GlyAlaIleCysTrrLysLeuValArgProLysArgLysHisGlyTrrLysSerAlaGlu 220
 Db 863 GGGCCTATCTGCACCAAGCTGGTGAGACCAAGCGGAGAACCGGAGCAACCACTCGGCGAG 922
 QY GluGluLeuAlaArgAlaGlyTrrPheLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
 Db 923 GAGAGAGTGGCAG 982
 QY IleGlyGlnGlyGluPheGlyAlaValLeuGlnGlyLysTrrLeuGlyGlnLysValAla 260
 Db 983 ATCGAAG 1042
 QY ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGlyThrAlaValMet 280
 Db 1043 GTGAAGAATATCAAGTGTGATGTACAGCCAGGCGCTTCTCGAGCAGAGAGCGGTCATG 1102
 QY ThrLysMetGlnHisGluAsnLeuValArgLeuGlnGlyAlaIleLeuHisGlnGlyLeu 300
 Db 1103 ACGAAGATGCACACGAGAACCTGTGCTCTCTGGGCGGTGATCTTCACCAAGAGGGCTG 1162
 QY TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgTrrArgGly 320
 Db 1163 TACATTTGTCATGAGCAGCAGTACAGAGCAACTGTGTAAGTCTTTCGCGAGACCGGGGT 1222
 QY ArgAlaLeuValAsnTrrAlaGlnLeuGlnPheSerLeuHisValAlaGlnGlyMet 340
 Db 1223 CGAGCCCTCTGGAACCGCCCTCAGCTCTGAGTTTCTCTGACGTCGCGGAGAGGGATG 1282
 QY GluTrrLeuGlnSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
 Db 1283 GAGTACTGTCGAGACCAAGAGCTGTGCACCGCACTGGCGCGCCGCAACATCTCTGTGTC 1342
 QY SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
 Db 1343 TCAGAGAGACCTGGTGGCCAGAGTACGAGACTTGGCTGGCGCAAGCCGAGCGGAGAGGG 1402
 QY LeuAspSerSerArgLeuProValLysTrrPrrAlaProGluAlaLeuLysHisGlyLys 400
 Db 1403 CTAGACTCAAGCCGCGTCCGCTCAAGTGAAGCGGCGCGAGGCTCTCAACACAGGG-- 1459

QY 401 PheThrSerLysSerAspValTrrPrrPheGlyValLeuLeuTrrGluAlaPheSerTyr 420
 Db 1460 TTCACCGCAGCAAGTCGATGTCTGAGATTTTGGGGTGTCTGCTCTGGAGAGCTTCTCATAT 1519
 QY 421 GlyArgAlaProTrrPrrLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
 Db 1520 GGAAGGCGCTCCGTACCTTAATAATGTCACTGAAGAAGAGTGTGCGAGAGAGAGAGAGAGAG 1579
 QY 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerSerCys 460
 Db 1580 TACCGCATGGAACCCCGGAGGGCTGTCAAGGCCCTGTGACGCTCATGACGAGCGTGC 1639
 QY 461 TrrPrrAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
 Db 1640 TGGGAGCAG 1698
 QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
 Db 1699 GAGCTACGACAGTGCAGAGTGGCCCGAGCTCGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAG 1757
 QY 501 SerProArgSerGlnGluPro 507
 Db 1758 TCGCCCGCAAGCAG 1778
 RESULT 14
 HSM802388
 LOCUS HSM802388 1804 bp mRNA linear PRI 18-FEB-2000
 DEFINITION Homo sapiens mRNA; cDNA DKFZp434N1212 (from clone DKFZp434N1212); partial cds.
 ACCESSION AL137754
 VERSION AL137754.1 GI:6808456
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1804)
 AUTHORS Duesterhoeft,A., Lauber,J., Mewes,H.W., Well,B. and Wiemann,S.
 TITLE Direct Submission
 JOURNAL Submitted (15-JUN-2000) MTPS, Am Kiofierspitz 18a, D-82152 Martinsried, GERMANY
 COMMENT
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de; sequenced by Olegen (Hilden/Germany) within the CDNA sequencing consortium of the German Genome Project.
 This clone (DKFZp434N1212) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.
 location/Qualifiers
 1..1804
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /map="19p13"
 /clone="DKFZp434N1212"
 /tissue_type="testis"
 /clone_lib="434 (synonym: htes3). Vector pSport1; host DH10B; sites NotI + SalI"
 /dev_stage="adult"
 1..1804
 /gene="DKFZp434N1212"
 <1..1574
 /gene="DKFZp434N1212"
 /note="protein-tyrosine kinase (EC 2.7.1.112) (Homo sapiens)"
 /product="hypothetical protein"
 /codon_start=3
 /protein_id="CAB70906.1"
 /db_xref="GI:6808457"
 /translation="GNGATRSKLCSDRLGVCVSRSLAEALIGAGAGRGSGAPPEVSPR

FLRANHPPVSAKMPTRMAEGTCTTCENTRPRKRGELAFKRGDVTILLACENKSM
 YVKKHTSGOGLLAAGLAREALASAPKLSLMPFMRKISGOEVLQDLPDEGLF
 LVRESARHFGDYVLCSEFGRDYIHYRVLHMDHIDEVAFKISLMDMEHESKXGA
 ICTKLVRPRKRGKTSABEELARAGMLNLJOLTGAOIGEEFPVLOGEYVNGKVA
 VNKICDVTAAQFLDETAFTKMOHENYVRLIGLHOGLYTVMEHVSNGNLYNFKRT
 RRRLAYNTRQLOQFSLHVAEEMETLESKULVHRLAARITLYSEDLYAKVSPGLAKA
 ERKGDSSRLPYKMTAPKALHKGFTSKSDVSGVLMEVSYGRAPPKSLKLSVS
 EAVEKGYRMEPEEGCPGVHVLMSSEWEAPARRPFRKLARELRLSAGAPASVS
 GODADGSTSRSQEP"
 1782
 polyA_site
 /gene="DKFZP43AN1212"

BASE COUNT 364 a 557 c 614 g 269 t
 ORIGIN

Alignment Scores:

Pred. No.: 1.6e-150 Length: 1804
 Score: 2544.00 Matches: 488
 Percent Similarity: 96.648 Conservative: 1
 Best Local Similarity: 96.448 Mismatches: 5
 Query Match: 95.254 Indels: 12
 DB: 9 Gaps: 1

US-09-977-260-2 (1-507) x HSM802388 (1-1804)

QY 2 AAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGluGlu 21
 DB 90 GCGGGGCGCGGC-----TCGGGGGCGCGCC 113
 QY 22 LeuProArgValSerProArgPheLeuArgAlaTrpHisProProArgValSerAlaArg 41
 DB 114 CCTGAGACAGGTAGCCCCCGCTTCCTCGAGGCTGGACCCCGCTCCGCTCAGCGCAGG 173
 QY 42 MetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGlnHisThrArg 61
 DB 174 ATGCCAAGAGCGCGCGGCGCCCGGACCCAGGTGATCACCAGCAATGCCAGACACCCGC 233
 QY 62 ProLysProGlyGlyLeuAlaPheArgGlyGlyAspValAlaThrIleLeuGlnAlaCys 81
 DB 234 CCCAAGCCAGGGAGCGTGGCTTCGCGAAGGGCGAGCTGGTCACCATCTCGAGAGGCTTC 293
 QY 82 GluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGlnGlyLeuLeuAla 101
 DB 294 GAGAACAGAGCTGTGACCGCTCAAGCCACACAGTGGACAGAGAGGGCTGCTGCGCA 353
 QY 102 AlaGlyAlaLeuArgGlyLysAlaLeuSerAlaAspProLysLeuSerLeuMetPro 121
 DB 354 GCTGGGCGCGCTCGGAGCGGAGCGCCCTCCGCGAGAGCCCAAGCTCAAGCTCATGCGG 413
 QY 122 TrpPheHisGlyLysIleSerGlyGlnGlnAlaValGlnGlnLeuGlnProProGluAsp 141
 DB 414 TGGTTCACAGGGAAGATCTCGGGCCAGAGGCTGTCCAGCAGCTGACGCTCCGAGGAT 473
 QY 142 GlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuGlyValSer 161
 DB 474 GGGCTGTTCTCTGTCGGGAGTCCCGCGCCACCCCGGAGCTAGCTCTGTGCGTGAAC 533
 QY 162 PheGlyArgAspValIleHisTrpArgValLeuHisArgAspGlyHisLeuThrIleAsp 181
 DB 534 TTGGCGCGGACGTCATCCATCCGCGGTGCGACCGGAGCGGCGACCTCATCATGAT 593
 QY 182 GluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTrpSerLysAspLysGly 201
 DB 594 GAGGCGGTTCTTCTGCAACCTCATGACATGGGAGGACATTCAGCAAGACCAAGAGGC 653
 QY 202 AlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGluGlu 221
 DB 654 GCTATCTGCACCAAGCTGTGAGACCAAGCGGAACAGGAGCAAGTGGCGCGAGGAG 713
 QY 222 GluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnIle 241
 DB 714 GAGCTGGCGCAGGCGGCGGTGTTACTGAACCTGCGACATTTGACATTGGGAGACACAGATC 773
 QY 242 GlyGluGlyGluPheGlyAlaValLeuGlnGlyGluTrpLeuGlyGlnLysValAlaVal 261

DB 774 GAGAGGGGAGCTTTGAGACTCTCTCGCAGGGGTGATACCTGGGCGCAAAAGGTGCGCG 833
 QY 262 LysAsnIleLysCysAspValIleHisArgAlaPheLeuAspGlyThrAlaValMetThr 281
 DB 834 AAGATATATCAAGTGTGATGTGACAGCCAGGCTTCTGTGACAGACAGGCGCTCATAGAC 893
 QY 282 LysMetGlnHisGlyLysLeuValArgLeuGlyValIleLeuHisGlnGlyLeuTrp 301
 DB 894 AAGATGCACACAGAGAACCTGTGGCGTCTCGTGGGCGTATCTGCACACAGGCGGTGAC 953
 QY 302 IleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGlyArg 321
 DB 954 ATTGTATGATGAGACAGTGTGAGCAAGGCAACCTGTGAACTTCTCGGCGCGCGGTGCA 1013
 QY 322 AlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlnGlyMetGlu 341
 DB 1014 GCCCTGTGTAACACCGCTGAGCTCTGCGAGTTTCTGTGACAGTGGCGGAGGCGATGAG 1073
 QY 342 TyrLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuValSer 361
 DB 1074 TACCTGAGAGAGCAAGAGCTGTGCACCGCAGCTGGCGCGCGCAACATCTGTGATCTCA 1133
 QY 362 GluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGlyLeu 381
 DB 1134 GAGACCTGTGTGCCAAGGTGACAGCACTTGGCTGGCCAAAGCGAGCGGAAGGGCTTA 1193
 QY 382 AspSerSerArgLeuProValLysTrpTrpAlaProGlyAlaLeuLysHisGlyLysPhe 401
 DB 1194 GACTCAAGCCGCGCTGCGCTCAAGTGGACGGCGCGCGAGGCTCTGAACACGGGAATTC 1253
 QY 402 ThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrpGly 421
 DB 1254 ACCAGCAAGTCCGAGTGTGGAGTTTGGGGTGTGCTCTGAGAGAGTCTTCATATATGGA 1313
 QY 422 ArgAlaProTrpTrpProLysMetSerLeuLysGluValSerGluAlaValGluLysGlyTrp 441
 DB 1314 CGGGCTCGGTACCTTAATATGTCAGTGAAGAGGTGTGCGAGGCGGTGGAAGAGGGTAC 1373
 QY 442 ArgMetGluProProGlyGlyCysProGlyProValHisValLeuMetSerSerCysTrp 461
 DB 1374 CGCATGGAACCCCGGAGGGCTGTCCAGGCGCCCGTGCACAGTCTCATAGACGCTGTGG 1433
 QY 462 GluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArgGlu 481
 DB 1434 GAGCGAGAGCCCGCGCGCGCCACCTTCCCAAACTGGCGGAAGCTGGCCCGGAG 1493
 QY 482 LeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThrSer 501
 DB 1494 CTACGCAAGTGCAGGTGCCCACCTCGTCTCAGGCGAGAGCGCGAGGCTCCACCTGG 1553
 QY 502 ProArgSerGlnGluPro 507
 DB 1554 CCGCGAAGCGAGAGGCC 1571

RESULT 15
 AK055395 2073 bp mRNA linear PRI 01-AUG-2002
 LOCUS Homo sapiens cDNA FLJ30833 fis, clone FEBRA2002027, highly similar
 DEFINITION to MEGAKAROCYTE-ASSOCIATED TYROSINE-PROTEIN KINASE (EC 2.7.1.112).
 ACCESSION AK055395.1 GI:16550112
 VERSION AK055395.1
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS
 Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,
 Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
 Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
 Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
 Wagaitsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,

Mon Aug 4 14:40:04 2003

us-09-977-260-2.rge

Page 18

|||||
Db 1857 GAGCC 1862

Search completed: August 1, 2003, 21:43:10
Job time : 6062 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:35:28 ; Search time 9490 Seconds

(without alignments)
8621.628 Million cell updates/sec

Title: US-09-977-260-1

Perfect score: 2000
Sequence: 1 ctgcctccaagtgtgcagc.....attcctaagactctaaanaa 2000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 2888711 segs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_ba: *
3: gb_in: *
4: gb_in: *
5: gb_in: *
6: gb_ov: *
7: gb_ov: *
8: gb_ov: *
9: gb_ov: *
10: gb_ov: *
11: gb_ov: *
12: gb_ov: *
13: gb_ov: *
14: gb_ov: *
15: gb_ov: *
16: gb_ov: *
17: gb_ov: *
18: gb_ov: *
19: gb_ov: *
20: gb_ov: *
21: gb_ov: *
22: gb_ov: *
23: gb_ov: *
24: gb_ov: *
25: gb_ov: *
26: gb_ov: *
27: gb_ov: *
28: gb_ov: *
29: gb_ov: *
30: gb_ov: *
31: gb_ov: *
32: gb_ov: *
33: gb_ov: *
34: gb_ov: *
35: gb_ov: *
36: gb_ov: *
37: gb_ov: *
38: gb_ov: *
39: gb_ov: *
40: gb_ov: *
41: gb_ov: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2000	100.0	2000	6	AR179662
2	2000	100.0	2000	6	AR274487
3	1971.2	98.6	2180	9	BC000114 Homo sapi
4	1971.2	98.6	2180	9	BC000114 Homo sapi
5	1963.4	98.2	1997	6	BD167710 Method of
6	1909.8	95.5	1987	6	AR084686
7	1909.8	95.5	1987	6	AR204704
8	1909.8	95.5	1987	6	AR053292
9	1909.8	95.5	1987	6	AR053292
10	1903.4	95.2	1942	6	E10345
11	1903.4	95.2	1942	6	HSMB02388
12	1669.2	83.5	1804	9	AK053395
13	1663	83.2	2073	9	AK053395
14	1651.4	82.6	1744	9	S71669
15	1519.4	76.0	1521	6	AR053291
16	1398	69.9	1398	6	AR053290
17	1377	68.8	1713	6	AX644057
18	1103.2	55.2	1734	10	MDSNRK
19	1071.4	53.6	1911	10	MDSNRK
20	1068.6	53.4	1651	10	MMU05210
21	1066	53.3	1656	10	S77473
22	1060	53.0	1838	10	RATBATK
23	738	36.9	738	6	AR053289
24	547.6	27.4	2187	9	AX330746
25	547.6	27.4	2187	9	HSCSRCKIN
26	547.6	27.4	2420	9	HSCYDCTK
27	537	26.9	1682	9	AY007162
28	521	26.1	1353	5	CHRSCKA
29	506	25.3	2125	10	REYTKI
30	496.4	24.8	2292	10	BC018394
31	496.4	24.8	2363	10	BC052006
32	491.6	24.6	1779	10	MMU05247
33	455	22.8	16389	9	AX644059
34	455	22.8	16389	9	AC005777
35	440	22.0	457	9	H006874503
36	372.2	18.6	1806	5	AF052430
37	281.6	14.1	2205	3	AK113892
38	272.4	13.6	1733	3	AB006558
39	236.2	11.8	245	9	H006874504
40	225	11.2	225	6	AR053288
41	220	11.0	233	9	H006874505
42	217.4	10.9	233	9	H006874502
43	216.4	10.8	1589	9	HS007236
44	215.2	10.8	1982	5	GGYRKA
45	213.6	10.7	3299	6	BD160548

ALIGNMENTS

RESULT 1
LOCUS AR179662
DEFINITION Sequence 1 from patent US 6326469.
ACCESSION AR179662
VERSION AR179662.1 GI:20221217
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2000)
AUTHORS Ullrich,A., Gishizky,M. and Sures,I.
TITLE Megakaryocytic protein tyrosine kinases
JOURNAL Patent: US 6326469-A 1 04-DEC-2001;
FEATURES Location/Qualifiers

Pred. No. is the number of results predicted by chance to have a

source 1. 2000
BASE COUNT 375 a 632 c 673 g 320 t
ORIGIN
Query Match 100.0%; Score 2000; DB 6; Length 2000;
Best Local Similarity 100.0%; Pred. No. 3.2e-289;
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CTCGCTCCAAAGTTGTGACGCGGACCGCTCGGGGTGTGACGCGGCTCGGGAGGCC 60
1 CTGGCTCCAAAGTTGTGACGCGGACCGCTCGGGGTGTGACGCGGCTCGGGAGGCC 60
Db
61 TCCTGG 120
61 TCCTGG 120
Db
121 AGGCTGGGTGAGTGGGACCGAGCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
121 AGGCTGGGTGAGTGGGACCGAGCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db
181 GGGCATTCCAGCGCTCCCGCAGCTGTGACCACTTGTCTGCTGCTGCTGCTGCTGCTG 240
181 GGGCATTCCAGCGCTCCCGCAGCTGTGACCACTTGTCTGCTGCTGCTGCTGCTGCTG 240
Db
241 TTTTCCCTGTGG 300
241 TTTTCCCTGTGG 300
Db
301 GCTGTGATTTGCTGAGAGAACTTCCCGGGGTGAGCCCGGCTTCTTCCGAGCGCTGG 360
301 GCTGTGATTTGCTGAGAGAACTTCCCGGGGTGAGCCCGGCTTCTTCCGAGCGCTGG 360
Db
361 CCCCTCCGCTGTGAGCGGAGTGCACAGAGGGGCTGGGCCCGGGCCGACCGAGTATCA 420
361 CCCCTCCGCTGTGAGCGGAGTGCACAGAGGGGCTGGGCCCGGGCCGACCGAGTATCA 420
Db
421 CCAAAATGCGAGCACACCCCGCCCAAGCGAGGGAGCTGGCTTCCGAAAGGGGAGCTGG 480
421 CCAAAATGCGAGCACACCCCGCCCAAGCGAGGGAGCTGGCTTCCGAAAGGGGAGCTGG 480
Db
481 TCACATCTGTGAGGCGCTGCGAGAACAAAGAGTGTGCTGCTGCTGCTGCTGCTGCTG 540
481 TCACATCTGTGAGGCGCTGCGAGAACAAAGAGTGTGCTGCTGCTGCTGCTGCTGCTG 540
Db
541 GACAGAGGGGGGCTGTGTGAGCTGTGGGGGCTGTGGGGGAGGAGGCGGCTTCCGAGACC 600
541 GACAGAGGGGGGCTGTGTGAGCTGTGGGGGCTGTGGGGGAGGAGGCGGCTTCCGAGACC 600
Db
601 CCAAGCTCAGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
601 CCAAGCTCAGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db
661 AGCTGAGCGCTCCCGAGGATGGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
661 AGCTGAGCGCTCCCGAGGATGGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db
721 ACTACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
721 ACTACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db
781 AGGGCCACCTCACAATGAGTGGGCGGTGTTCTTCTGCAACCTCATGAGCATGTGTGAGC 840
781 AGGGCCACCTCACAATGAGTGGGCGGTGTTCTTCTGCAACCTCATGAGCATGTGTGAGC 840
Db
841 ATTACAGCAAGAGCAAGGGGCTATCTGACCAAGCTGTGTGAGACCAAGGAGGAGGAGC 900
841 ATTACAGCAAGAGCAAGGGGCTATCTGACCAAGCTGTGTGAGACCAAGGAGGAGGAGC 900
Db
901 GGAACCAAGTGGCGGAGAGGAGCTGTGCCAGGGGCGGGCTGTGTTACTGAACCTGACGATT 960
901 GGAACCAAGTGGCGGAGAGGAGCTGTGCCAGGGGCGGGCTGTGTTACTGAACCTGACGATT 960
Db

961 TGACATTGGAGACACAGATGCGAGAGGAGATTGTGAGCTGTCTGTCAGAGGTAGTACC 1020
961 TGACATTGGAGACACAGATGCGAGAGGAGATTGTGAGCTGTCTGTCAGAGGTAGTACC 1020
Db
1021 TGGGGCAAAAGTGGCCGTGAAGATATCAAGTGTATGTGACAGCCCAAGCCTTCTGTGG 1080
1021 TGGGGCAAAAGTGGCCGTGAAGATATCAAGTGTATGTGACAGCCCAAGCCTTCTGTGG 1080
Db
1081 ACAGAGCGCCGTATGACAGAAATGACACAGCAAGCACTGTGTGCTGCTGCTGCTGCTG 1140
1081 ACAGAGCGCCGTATGACAGAAATGACACAGCAAGCACTGTGTGCTGCTGCTGCTGCTG 1140
Db
1141 TCCTGACCAAGGGGCTGTATCTGATGTGATGAGCAGCTGAGCAAGGGCAACCTGTGAACT 1200
1141 TCCTGACCAAGGGGCTGTATCTGATGTGATGAGCAGCTGAGCAAGGGCAACCTGTGAACT 1200
Db
1201 TTTCTGGGAGCCCGGGGGTGGAGCCCTGTGTAACCCGCTGACGCTGCTGCTGCTGCTGCTG 1260
1201 TTTCTGGGAGCCCGGGGGTGGAGCCCTGTGTAACCCGCTGACGCTGCTGCTGCTGCTGCTG 1260
Db
1261 ACCTGGCCGAGGCGATGAGTACCTGTGAGAGCAAGAACTTGTGCAACCGGACCTGTGGCC 1320
1261 ACCTGGCCGAGGCGATGAGTACCTGTGAGAGCAAGAACTTGTGCAACCGGACCTGTGGCC 1320
Db
1321 CCGGCAACATCTGTGTCTGAGAGGAGCTGTGGCCCAAGGTACAGCTTGTGGCTGTGGCCA 1380
1321 CCGGCAACATCTGTGTCTGAGAGGAGCTGTGGCCCAAGGTACAGCTTGTGGCTGTGGCCA 1380
Db
1381 AAGCCGAGCGAAGGGGCTGAGCTCAACAGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 1440
1381 AAGCCGAGCGAAGGGGCTGAGCTCAACAGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 1440
Db
1441 CTCTCAAAACCGGGAAGTTTACACAGCAAGTGTGATTTTGTGGGGTCTGCTGCTGCTG 1500
1441 CTCTCAAAACCGGGAAGTTTACACAGCAAGTGTGATTTTGTGGGGTCTGCTGCTGCTG 1500
Db
1501 GGGAGGCTTCTTATATGAGCGGGCTTCCGTAACCTTAAAGTGTACTGAAGAGAGTGTGCG 1560
1501 GGGAGGCTTCTTATATGAGCGGGCTTCCGTAACCTTAAAGTGTACTGAAGAGAGTGTGCG 1560
Db
1561 AGGCGGTGAGAGAGGGGTACCGCATGGAACCCCGGAGGGCTTCCAGGGCCCGCTGACAG 1620
1561 AGGCGGTGAGAGAGGGGTACCGCATGGAACCCCGGAGGGCTTCCAGGGCCCGCTGACAG 1620
Db
1621 TCCTCATGAGCACTGTGTGAGGAGGAGCCCGCGGCAACCTTCCGCAAACTGG 1680
1621 TCCTCATGAGCACTGTGTGAGGAGGAGCCCGCGGCAACCTTCCGCAAACTGG 1680
Db
1681 CCGAGAAAGTGGCGCCCGGAGAGCTACGAGTGCAGAGTGCCTCCAGCTTCCGCAAACTGG 1740
1681 CCGAGAAAGTGGCGCCCGGAGAGCTACGAGTGCAGAGTGCCTCCAGCTTCCGCAAACTGG 1740
Db
1741 ACGCCAGAGGCTCACCTGCGCCCGAAGCGAGAGGCTGTGACCCACCGGAGGAGGAGCCT 1800
1741 ACGCCAGAGGCTCACCTGCGCCCGAAGCGAGAGGCTGTGACCCACCGGAGGAGGAGCCT 1800
Db
1801 TGGCCCCAGAGACCGAGAGAGTGTGAGAGTGTGGGGCTGTGAGGAGGAGGAGGAGGAG 1860
1801 TGGCCCCAGAGACCGAGAGAGTGTGAGAGTGTGGGGCTGTGAGGAGGAGGAGGAGGAG 1860
Db
1861 AGGGTCCAGGGGGGCAAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
1861 AGGGTCCAGGGGGGCAAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
Db
1921 GGGCTGTGG 1980
1921 GGGCTGTGG 1980
Db
1981 ATTCTAAGAGCTCTAAAAA 2000
1981 ATTCTAAGAGCTCTAAAAA 2000
Db

RESULT 2
LOCUS AR274487 2000 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1 from patent US 6506578.
ACCESSION AR274487
VERSION AR274487.1 GI:29706942
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2000)
AUTHORS Ullrich,A., Gishizky,M. and Sures,I.G.
TITLE Nucleotide encoding megakaryocytic protein tyrosine kinases
JOURNAL Patent: US 6506578-A 1 14-JAN-2003;
FEATURES
source location/Qualifiers
BASE COUNT 375 a 632 c 673 g 320 t
ORIGIN
Query Match 100.0%; Score 2000; DB 6; Length 2000;
Best Local Similarity 100.0%; Pred. No. 3.2e-289;
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGCTCCAGATTGTGACGCCGGACCGCTCTGGGGTGTGACAGCCGGCTCGGGAGGCC 60
DB 1 CTCGCTCCAGATTGTGACGCCGGACCGCTCTGGGGTGTGACAGCCGGCTCGGGAGGCC 60
QY 61 TCTTGG 120
DB 61 TCTTGG 120
QY 121 AGGCTGGTCCAGTGGGACCCAGCTCCCTACCTCTGTGGCCAGCCGGCGCTGGGGA 180
DB 121 AGGCTGGTCCAGTGGGACCCAGCTCCCTACCTCTGTGGCCAGCCGGCGCTGGGGA 180
QY 181 GGCATTTCCAGCGCTCCCGGACTGTGACCACTTGTCTAGTGTGCTTACCTGCTCAG 240
DB 181 GGCATTTCCAGCGCTCCCGGACTGTGACCACTTGTCTAGTGTGCTTACCTGCTCAG 240
QY 241 TTTCCCTCTGG 300
DB 241 TTTCCCTCTGG 300
QY 301 GCTGATTTGCTGAGGAATTTCCCGGGGTGAGCCCGGCTTCTCCGAGCGTGGCAC 360
DB 301 GCTGATTTGCTGAGGAATTTCCCGGGGTGAGCCCGGCTTCTCCGAGCGTGGCAC 360
QY 361 CCCCTCCGCTGCACCGAGATGCCAGAGGCGCTGGGCCCGCGGACCCAGTGTATCA 420
DB 361 CCCCTCCGCTGCACCGAGATGCCAGAGGCGCTGGGCCCGCGGACCCAGTGTATCA 420
QY 421 CCAATTCGAGACACCCGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGCGACGTGG 480
DB 421 CCAATTCGAGACACCCGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGCGACGTGG 480
QY 481 TCACCATCTGGAGGCTGCGAACAAGATGTGTACCGGCTCAAGACCAACACACAGG 540
DB 481 TCACCATCTGGAGGCTGCGAACAAGATGTGTACCGGCTCAAGACCAACACACAGG 540
QY 541 GACAGAGAGGGGCTGCTGGACAGTGGGGGCTGCGGGAGCGGGAGGCCCTTCTCCGAG 600
DB 541 GACAGAGAGGGGCTGCTGGACAGTGGGGGCTGCGGGAGCGGGAGGCCCTTCTCCGAG 600
QY 601 CCAAGCTACGCTTATGCGGTGTTCACAGGGAATATCTGGGCCAGAGAGGTGTCCAGC 660
DB 601 CCAAGCTACGCTTATGCGGTGTTCACAGGGAATATCTGGGCCAGAGAGGTGTCCAGC 660
QY 661 AGCTCAGACCTCCAGAGAGGAGTGTCTGTTCTGGGAGTCCGCGGCGACCCCGGGG 720
DB 661 AGCTCAGACCTCCAGAGAGGAGTGTCTGTTCTGGGAGTCCGCGGCGACCCCGGGG 720
QY 721 ACTAGTCCTGTGCTGAGCTTTGGCCGCGAGCGTATCCACTACCGCGTGTGACCGCG 780

DB 721 ACTAGTCCTGTGCTGAGCTTTGGCCGCGAGCGTATCCACTACCGCGTGTGACCGCG 780
QY 781 ACGCCACCTCAACAATTCATGAGCGCGTGTCTTCTGCAACCTATGACATGTGAGC 840
DB 781 ACGCCACCTCAACAATTCATGAGCGCGTGTCTTCTGCAACCTATGACATGTGAGC 840
QY 841 ATTACAGCAGGAGGAGGCGCTATCTGACCAACCTGTGAGGAGGAGGAGGAGGAGG 900
DB 841 ATTACAGCAGGAGGAGGCGCTATCTGACCAACCTGTGAGGAGGAGGAGGAGGAGG 900
QY 901 GGAGCAAGTCGGCCGAGAGAGAGTGGCCAGGGCGGGTGTGTACTGAACTGCAGATT 960
DB 901 GGAGCAAGTCGGCCGAGAGAGAGTGGCCAGGGCGGGTGTGTGTACTGAACTGCAGATT 960
QY 961 TGACATTGGAGACACAGATCGAGAGGAGAGTGTGAGTGTCTGACGGGTGAGTACC 1020
DB 961 TGACATTGGAGACACAGATCGAGAGGAGAGTGTGAGTGTCTGACGGGTGAGTACC 1020
QY 1021 TGGGCAAAAGGTGGCCGCGAAGATATCAAGTGTGATGTGACAGCCAGGCGTTCG 1080
DB 1021 TGGGCAAAAGGTGGCCGCGAAGATATCAAGTGTGATGTGACAGCCAGGCGTTCG 1080
QY 1081 ACGAGACGCGCTCATGACGAAGATGCAACACGAAACCTGGTCTCTCGGGCGTGA 1140
DB 1081 ACGAGACGCGCTCATGACGAAGATGCAACACGAAACCTGGTCTCTCGGGCGTGA 1140
QY 1141 TCTCTACACAGGGGCTGTACATTTGTATGAGACACGTAGACAGGCAACCTGGTGA 1200
DB 1141 TCTCTACACAGGGGCTGTACATTTGTATGAGACACGTAGACAGGCAACCTGGTGA 1200
QY 1201 TTTCTGGAGCCGGGGGTGAGAGCCCTGTAACACACGCTCAGTCTCTGAGTTTCTG 1260
DB 1201 TTTCTGGAGCCGGGGGTGAGAGCCCTGTAACACACGCTCAGTCTCTGAGTTTCTG 1260
QY 1261 ACGTGGCCGAGGGATGAGATACCTGAGAGCAAGAACTTGTGACCGCGACCTGAGC 1320
DB 1261 ACGTGGCCGAGGGATGAGATACCTGAGAGCAAGAACTTGTGACCGCGACCTGAGC 1320
QY 1321 CCCGCAATCTGTGTCTCAGAGACCTGTGTGAGGCAAGTCAAGCAGCATTTGGCTG 1380
DB 1321 CCCGCAATCTGTGTGTCTCAGAGACCTGTGTGAGGCAAGTCAAGCAGCATTTGGCTG 1380
QY 1381 AAGCGAGCGGAGGGGTGAGATCAACACCGGCTGCTGAAGTGAAGGAGGCGCCGAG 1440
DB 1381 AAGCGAGCGGAGGGGTGAGATCAACACCGGCTGCTGAAGTGAAGGAGGCGCCGAG 1440
QY 1441 CTCTCAACACAGGGAGTTTCAACAGCAAGTGGATGTCTGAGTTTGGGGTGTGCTCT 1500
DB 1441 CTCTCAACACAGGGAGTTTCAACAGCAAGTGGATGTCTGAGTTTGGGGTGTGCTCT 1500
QY 1501 GGGAGGTTTCTCATATGAGAGGGGCTCCGTACCCCTAAATGTCACTGAAAGGAGT 1560
DB 1501 GGGAGGTTTCTCATATGAGAGGGGCTCCGTACCCCTAAATGTCACTGAAAGGAGT 1560
QY 1561 AGCGGTGAGAGGGGTACCGCATGGAACCCCGGAGGGGCTGACAGGCCCGGTGACG 1620
DB 1561 AGCGGTGAGAGGGGTACCGCATGGAACCCCGGAGGGGCTGACAGGCCCGGTGACG 1620
QY 1621 TCTTCATAGAGAGTGTGAGAGGAGCCCGCGGCGACACCTTCCGCAAACTGG 1680
DB 1621 TCTTCATAGAGAGTGTGAGAGGAGCCCGCGGCGACACCTTCCGCAAACTGG 1680
QY 1681 CCGAGAACTGCGCCGAGAGTACGCAATGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1740
DB 1681 CCGAGAACTGCGCCGAGAGTACGCAATGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1740
QY 1741 ACGCGAGGCTCCACTCGCCCGCAAGCAGAGACCTTGAACCCACCGGTGGGGCCT 1800
DB 1741 ACGCGAGGCTCCACTCGCCCGCAAGCAGAGACCTTGAACCCACCGGTGGGGCCT 1800
QY 1801 TGGCCCAAGAGACCGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 1860

	Db	1801	TGGCCCCAGAGACCGAGAGAGTGGAGAGTGCGGCGTGCGGGCCACTGCACCAGGCCCAAGG	1860
OY	1861	AGGCTTCAGCGCGGGCAGATCATCTCTCTGTGTGCCCACAGCAGGGGCTGTGCCCACGTAGGG	1920	
Db	1861	AGGGTCCAGGGGGGAATCATCTCTCTGTGTGCCCACAGCAGGGGCTGTGCCCACGTAGGG	1920	
OY	1921	GCGTCTGGGGCGCGCGTGAGCACCCGACAGCTGGGGAAGATGATCGCCGATTAAAGACGG	1980	
Db	1921	GGCTCTGGGGCGCGCGTGAGCACCCGACAGCTGGGGAAGATGATCGCCGATTAAAGACGG	1980	
OY	1981	ATTCTAAGAGACTTAATAAAA	2000	
Db	1981	ATTCTAAGAGACTTAATAAAA	2000	
RESULT 3				
BC000114				
LOCUS		2180 bp	mRNA	linear PRI 12-JUL-2001
DEFINITION		Homo sapiens, megakaryocyte-associated tyrosine kinase, clone		
ACCESSION		MGC:17108 IMAGE:3350365,	mRNA, complete cds.	
VERSION		BC000114.1	GI:12652728	
KEYWORDS		MGC.		
SOURCE		Homo sapiens (human)		
ORGANISM		Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE		1 (bases 1 to 2180)		
JOURNAL		Strausberg, R. Direct Submission Submitted (03-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK		NIH-MGC Project URL: http://mgc.ncl.nih.gov		
COMMENT		Contact: MGC help desk Email: gcgpb@rmail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) Sequence by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@cgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chin, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Carolee McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schel, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.		
FEATURES		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 6 Row: 1 Column: 3 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505108. Location/Oualifiers 1..2180 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="LocusID:4145" /db_xref="taxon:9606" /clone="MGC:1708 IMAGE:3350365" /tissue_type="Eye, retinoblastoma" /clone_lib="NH_MGC_16" /lab_host="DH10B-R" /note="Vector: pOTB7" 400..1923 /codon_start=1 /product="megakaryocyte-associated tyrosine kinase" /protein_id="AAH00114.1" /db_xref="GI:12652729"		
CDS				

Query Match	Best Local Similarity	Score	DB	Length
Matches 1996: Conservative	99.7%	99.7%	6.3e-285	Indels 3; Gaps 2;
BASE COUNT	436 a	689 c	724 g	331 t
ORIGIN				
1	CTCGCTCCAAATTGTGACCGCGGACCCCTCGGGGTGTGCAGCGCGCTCGGAGCGCC	60		
142	CTCGCTCCAAATTGTGACCGCGGACCCCTCGGGGTGTGCAGCGCGCTCGGAGCGCC	201		
61	TCCTGGGGGGGGGGGGGG--GGGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	118		
202	TCCTGG	261		
119	CCAGGCTGGGTCAGTGGACCCAGCTCCCACTCCCTGTCGACGCGCGCTGGCTGG	178		
262	CCAGGCTGGGTCAGTGGACCCAGCTCCCACTCCCTGTCGACGCGCGCTGGCTGG	321		
179	CAGGCGAATCCAGCGTCCCGAGCTGTACCACTTGTCACTGTGACCTCTCACTGCTC	238		
322	CAGGCGAATCCAGCGTCCCGAGCTGTACCACTTGTCACTGTGACCTCTCACTGCTC	381		
239	AGTTTCCCTGGGGGGGGGATGGCGGGGCGAGGCTCTGTGTTTCTTGGCGGGCAATTCA	298		
382	AGTTTCCCTGGGGGGGGGATGGCGGGGCGAGGCTCTGTGTTTCTTGGCGGGCAATTCA	440		
299	CGGCTGTATCTGCTGAGGAATTTCCCGGGGTGAGCCCCGGCTTCTCCGAGCTGGCA	358		
441	CGGCTGTATCTGCTGAGGAATTTCCCGGGGTGAGCCCCGGCTTCTCCGAGCTGGCA	500		
359	CCCCCTCCGCTCAGCGAGGATGCCAAGAGAGGCGTGGGGCCCGGGCACCAAGTAT	418		
501	CCCCCTCCGCTCAGCGAGGATGCCAAGAGAGGCGTGGGGCCCGGGCACCAAGTAT	560		
419	CACCAATGAGACACACCGCCCCCAAGGAGAGCTGGCTTCCGCAAGGGCGAGCT	478		
561	CACCAATGAGACACACCGCCCCCAAGGAGAGCTGGCTTCCGCAAGGGCGAGCT	620		
479	GGTCAACATCTGTGAGGCGTGGAGAACAAAGAGTGAACGGCGTCAAGCACACACAG	538		
621	GGTCAACATCTGTGAGGCGTGGAGAACAAAGAGTGAACGGCGTCAAGCACACACAG	680		
539	TGACAGAGAGGGGCTGTGAGAGTGGGGGCGCTGGGGAGGGGAGGCGCTCTCCGAGA	598		
681	TGACAGAGAGGGGCTGTGAGAGTGGGGGCGCTGGGGAGGGGAGGCGCTCTCCGAGA	740		
599	CCCCAAGCTCAGCTCATGCCGTGGTTCCAGCGGAAGATCTGGGGCCAGAGGCTGTCCA	658		
741	CCCCAAGCTCAGCTCATGCCGTGGTTCCAGCGGAAGATCTGGGGCCAGAGGCTGTCCA	800		
659	GCAGCTGAGCTCCCGAGAGTGGGCTTCTGTGTGGGAGTTCGCGCGCCACCCCGG	718		
801	GCAGCTGAGCTCCCGAGAGTGGGCTTCTGTGTGGGAGTTCGCGCGCCACCCCGG	860		
719	CGACTAGCTCTGTGGCGTGAAGCTTTGGCGCGGAGCTATCCATCAACCGCGTGCACCG	778		
861	CGACTAGCTCTGTGGCGTGAAGCTTTGGCGCGGAGCTATCCATCAACCGCGTGCACCG	920		
779	CGAGGGCCACTCAATCGATGAGGGCGTTCCTTCGCAACCTCATGAGCATGAGGAGA	838		
921	CGAGGGCCACTCAATCGATGAGGGCGTTCCTTCGCAACCTCATGAGCATGAGGAGA	980		
839	GCATTACGACAGGACAAAGGCGCTATCTGACCAAGCTGTGTGAGACCAAGCGGAACA	898		

```

|||||
Db 981 GCATTACAGCAAGGAGGAGGAGGCTATCTGCACCAAGCTGGTGAACACCAAGGAGGAGCA 1040
OY 899 CGGAGCAAGTCTGGCCGAGAGAGAGCTGGCCAGGGGGGCTGGTTACTGAACCTGCACCA 958
Db 1041 CGGGACCAAGTCTGGCCGAGAGAGAGCTGGCCAGGGGGGCTGGTTACTGAACCTGCACCA 1100
OY 959 TTTGACATTTGGAGACAGATCGGAGAGAGAGTTTGGAGTGTCTGCTGCAAGGGTGA 1018
Db 1101 TTTGACATTTGGAGACAGATCGGAGAGAGAGTTTGGAGTGTCTGCTGCAAGGGTGA 1160
OY 1019 CCTGGGGCAAAAAGGTGGCCGTGAGAAATATCAAGTGTATGTAGACAGCCAGGCTTCTCT 1078
Db 1161 CCTGGGGCAAAAAGGTGGCCGTGAGAAATATCAAGTGTATGTAGACAGCCAGGCTTCTCT 1220
OY 1079 GGAGAGAGAGCGCCGCTCATAGAGAGATGCAACACAGAGAACTGGTGGCTTCTCTGGCGT 1138
Db 1221 GGAGAGAGAGCGCCGCTCATAGAGAGATGCAACACAGAGAACTGGTGGCTTCTCTGGCGT 1280
OY 1139 GATCCTGACACAGGGGCTGTACATGTGTATGGAGACGTGAGCAAGGGCAACCTGTGTA 1198
Db 1281 GATCCTGACACAGGGGCTGTACATGTGTATGGAGACGTGAGCAAGGGCAACCTGTGTA 1340
OY 1199 CTTTCTGCGGAGCCCGGGGTGCAAGCCCTGCGTAACACGCTCAGCTCTGCAATTTTCTCT 1258
Db 1341 CTTTCTGCGGAGCCCGGGGTGCAAGCCCTGCGTAACACGCTCAGCTCTGCAATTTTCTCT 1400
OY 1259 GCAGCTGGCCGAGGGGCTAGAGTACTGAGAGAGCAAGAGCTGTGCAACCGGAGACTGGC 1318
Db 1401 GCAGCTGGCCGAGGGGCTAGAGTACTGAGAGAGCAAGAGCTGTGCAACCGGAGACTGGC 1460
OY 1319 GCAGCTGGCCGAGGGGCTAGAGTACTGAGAGAGCAAGAGCTGTGCAACCGGAGACTGGC 1378
Db 1461 GCAGCTGGCCGAGGGGCTAGAGTACTGAGAGAGCAAGAGCTGTGCAACCGGAGACTGGC 1520
OY 1379 CAAAGCCGAGCGGAGGAGGCTAGACTCAAGCCGCTGCCCTCAAGTGAGAGCGGCCGCA 1438
Db 1521 CAAAGCCGAGCGGAGGAGGCTAGACTCAAGCCGCTGCCCTCAAGTGAGAGCGGCCGCA 1580
OY 1439 GGCCTGCAAAACGAGGAGGAGTTCACAGAGAGCGAGTGTCTGAGAGTGTGGGGTCTCTGT 1498
Db 1581 GGCCTGCAAAACGAGGAGGAGTTCACAGAGAGCGAGTGTCTGAGAGTGTGGGGTCTCTGT 1640
OY 1499 CTGGGAGGCTCTCTCATATGAGAGGGGCTCCGTACCCCTAAATGTACTGAAAGAGTGTCT 1558
Db 1641 CTGGGAGGCTCTCTCATATGAGAGGGGCTCCGTACCCCTAAATGTACTGAAAGAGTGTCT 1700
OY 1559 GGAGGCGCTGAGAGAGGGGTACCGCATGAGAACCCCCCGAGGGGCTGTCCAGGCCCGTGCA 1618
Db 1701 GGAGGCGCTGAGAGAGGGGTACCGCATGAGAACCCCCCGAGGGGCTGTCCAGGCCCGTGCA 1760
OY 1619 GCTGCTCAATGAGACAGTGTGTGGAGAGGAGAGCCCGCCGCGGCAACCTTCCGGAANT 1678
Db 1761 GCTGCTCAATGAGACAGTGTGTGGAGAGGAGAGCCCGCCGCGGCAACCTTCCGGAANT 1820
OY 1679 GGCCGAGAGAGTGTGGCCCGGAGACTACGAGTGCAGAGTGGCCCAAGCTCCGCTCAGGGCA 1738
Db 1821 GGCCGAGAGAGTGTGGCCCGGAGACTACGAGTGCAGAGTGGCCCAAGCTCCGCTCAGGGCA 1880
OY 1739 GGAGGCGGAGGCTCCACTCGCCCGGAGAGCCAGAGAGCCCTGACCCACCCGCGTGGGGCC 1798
Db 1881 GGAGGCGGAGGCTCCACTCGCCCGGAGAGCCCTGACCCACCCGCGTGGGGCC 1940
OY 1799 CTTGGCCCCAGAGAGCCGAGAGAGTGGAGAGTGGCGGCTGGGGGCACTGACCAAGGCCCA 1858
Db 1941 CTTGGCCCCAGAGAGCCGAGAGAGTGGAGAGTGGCGGCTGGGGGCACTGACCAAGGCCCA 2000
OY 1859 GGAGGCTCCAGGCGGAGCAATCATCTCTCTGCTGCCACACAGCAGAGGGGCTGGCCCAAGTGA 1918
Db 2001 GGAGGCTCCAGGCGGAGCAATCATCTCTCTGCTGCCACACAGCAGAGGGGCTGGCCCAAGTGA 2060
OY 1919 GGGGCTCTGGCGGCGCCGTGACACCCAGACCTGCGAAGATGATCGGCCGATAAAGAC 1978

```

```

Db 2061 GGGGCTCTGGGCGCCCGCTGGACACCCGAGACTGTCGAAGATGATCCGCCGATAAAGAC 2120
OY 1979 GGATTTCTAAGCACTCTATAAAAA 2000
Db 2121 GGATTTCTAAGCACTATAAAAA 2142

RESULT 4
BC003109
LOCUS BC003109 2180 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, megakaryocyte-associated tyrosine kinase, clone
ACCESION MGC:2101 IMAGE:3357648, mRNA, complete cds.
VERSION BC003109
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2180)
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (13-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contract: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ruben Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LTLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Lee, Victor Ling, Carrie Matheson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Payvaneh Speedi, Jacqueline
Schein, Diane Smalls, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsal, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LTLNL at: http://image.lnl.nih.gov
Series: IRAL Plate: 6 Row: e Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4505108.
Location/Qualifiers
1. 2180
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="locusID:4145"
/db_xref="taxon:9606"
/clone="MGC:2101 IMAGE:3357648"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH_MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
400..1923
/codon_start=1
/product="megakaryocyte-associated tyrosine kinase"
/protein_id="AAH03109.1"
/db_xref="gi:13111883"
/translation="MAGRSLSVSWRAFHGCSAELPRVSPRELRAMHPYPVSAEMPT
RRMAGTQICITRCEHTRPKPEGLARFKDDVYVILACENKSVRYKHHTSGEGLLAA
GALREKALSDPKSLIMPEWFGKISGCAVOQLPPDEGLTVEESARHPDEYLVCA
SGRDIYHYRVLRHGH.LIDEAVFCQ.LIMDEYHSKDKAICTRLVPRKHGKFS
AEEELARGMNLNLOHTLGNIGSGEGGAYVCGEYLCKQAVKIKCDYVNOARFLE
TAVTRKMHENLVRLGVILHOGELTYMEHYSKGLVNFVLRTRGALVYNTAQLDPSL
HYADEMEYLESKLVLRDLAANILVSDVLAQVDFGLAERAGGLDSSRLPVWTA
PEALKHGFETSKSDVWSFGLMEVFSYGRAPVPMKSLKEVSEAVEKGYRMEPPGCP

```

BASE COUNT 436 a 689 c 724 g 331 t
ORIGIN

Query Match 98.6%; Score 1971.2; DB 9; Length 2180;
Best Local Similarity 99.7%; Pred. No. 6.3e-285;
Matches 1996; Conservative 0; Mismatches 3; Indels 3; Gaps 2;

1 CTCGCTCCAAAGTTGTGCAAGCCGGGAGCCGCTCGGGGTGTGAGCCGGCTCGCGAGGCC 60
Db CTCCTCCAAAGTTGTGCAAGCCGGGAGCCGCTCGGGGTGTGAGCCGGCTCGCGAGGCC 201
Oy TCCGCGGG 118
Db TCCGCGGG 261
Oy TCCGCGGG 178
Db TCCGCGGG 321
Oy TCCGCGGG 238
Db TCCGCGGG 381
Oy TCCGCGGG 298
Db TCCGCGGG 440
Oy TCCGCGGG 358
Db TCCGCGGG 500
Oy TCCGCGGG 418
Db TCCGCGGG 560
Oy TCCGCGGG 478
Db TCCGCGGG 620
Oy TCCGCGGG 538
Db TCCGCGGG 680
Oy TCCGCGGG 598
Db TCCGCGGG 740
Oy TCCGCGGG 658
Db TCCGCGGG 800
Oy TCCGCGGG 718
Db TCCGCGGG 860
Oy TCCGCGGG 778
Db TCCGCGGG 920
Oy TCCGCGGG 838
Db TCCGCGGG 980
Oy TCCGCGGG 898
Db TCCGCGGG 1040
Oy TCCGCGGG 958
Db TCCGCGGG 1100
Oy TCCGCGGG 1018
Db TCCGCGGG

1101 TTTGACATTTGGAGACACAGATCGAGAGGAGGAGTTTGGAGCTCTCTGACAGGCTGAGTA 1160
Oy CCTGGGGCAAAAAGGTGGCCGCTGAAGATATCAAGTGTGATGTGACAGCCAGGCTTCT 1078
Db CCTGGGGCAAAAAGGTGGCCGCTGAAGATATCAAGTGTGATGTGACAGCCAGGCTTCT 1220
Oy GGACGAGAGCGCCGCTCATGAGAGATGCAACACAGAGAACCTGTGTGCTCTCTGAGGCGT 1138
Db GGACGAGAGCGCCGCTCATGAGAGATGCAACACAGAGAACCTGTGTGCTCTCTGAGGCGT 1280
Oy GATCCTGACACAGGGGGCTGTATCATTTGATGTGAGACAGTGTGAGAGAGGCAACCTGTGTAA 1198
Db GATCCTGACACAGGGGGCTGTATCATTTGATGTGAGACAGTGTGAGAGAGGCAACCTGTGTAA 1340
Oy CTTTCTGCGGAGCCCGGGGGCTGAGCCCTCGTGAACACCCGCTCAGCTCTCTGAGTTTCTCT 1258
Db CTTTCTGCGGAGCCCGGGGGCTGAGCCCTCGTGAACACCCGCTCAGCTCTCTGAGTTTCTCT 1400
Oy GCAGGTGGCCGAGGCGATGAGTACCTGTGAGAGCAGAGAACCTGTGTGACCGGCACTGCGC 1318
Db GCAGGTGGCCGAGGCGATGAGTACCTGTGAGAGCAGAGAACCTGTGTGACCGGCACTGCGC 1460
Oy CGCCCGCAACATCTGTGTCTAGAGAGACCTGTGTGCGCAAGGTCAAGCACTTTGGCTGTGC 1520
Db CGCCCGCAACATCTGTGTCTAGAGAGACCTGTGTGCGCAAGGTCAAGCACTTTGGCTGTGC 1578
Oy CAAAGCCGAGCGAGAGGGGCTAGACTCAACCCGCTGCTGCAAGTGAAGGAGGCGCCGA 1438
Db CAAAGCCGAGCGAGAGGGGCTAGACTCAACCCGCTGCTGCAAGTGAAGGAGGCGCCGA 1580
Oy GCGCTCTCAAAACAGGGAGTTTCAACAGCAAGTGGATGTGTGAGTTTGGGGTGTGCT 1498
Db GCGCTCTCAAAACAGGGAGTTTCAACAGCAAGTGGATGTGTGAGTTTGGGGTGTGCT 1640
Oy CTGGGAGGCTCTCATATGAGAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1558
Db CTGGGAGGCTCTCATATGAGAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1700
Oy GGAAGGCGGTGAGAGAGGGGTACCGCATGTGAACCCCGGAGGCTGTCTCAAGGCGCCGTGA 1618
Db GGAAGGCGGTGAGAGAGGGGTACCGCATGTGAACCCCGGAGGCTGTCTCAAGGCGCCGTGA 1760
Oy GGTCTCTATAGACAGCTGCTGTGAGAGGAGAGCCCGGCGGCGGCACTTCCGCAAACT 1678
Db GGTCTCTATAGACAGCTGCTGTGAGAGGAGAGCCCGGCGGCGGCACTTCCGCAAACT 1820
Oy GGCCTGAGAACTGTGGCCCGGAGAGCTAGCGAGTGTGAGGCGCCGAGGCTGCTCAAGGCA 1738
Db GGCCTGAGAACTGTGGCCCGGAGAGCTAGCGAGTGTGAGGCGCCGAGGCTGCTCAAGGCA 1880
Oy GGAAGGCGGAGGCTTCACTGAGCCGAGAGCCGAGAGCCCTGACCCGAGGCTGTGAGGCA 1798
Db GGAAGGCGGAGGCTTCACTGAGCCGAGAGCCGAGAGCCCTGACCCGAGGCTGTGAGGCA 1940
Oy GGTGGCCCGGAGAGAGGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1858
Db GGTGGCCCGGAGAGAGGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 2000
Oy GGAAGGCTGAGGAGGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1918
Db GGAAGGCTGAGGAGGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 2060
Oy GGGGCTGTGGGCGGCGGTGAGACACCCGAGACCTGTGGAAGATGATGCGCCGTAAGAAC 1978
Db GGGGCTGTGGGCGGCGGTGAGACACCCGAGACCTGTGGAAGATGATGCGCCGTAAGAAC 2120
Oy GATTTCTAAGAGACTTAAGAA 2000
Db GATTTCTAAGAGACTTAAGAA 2142

RESULT 5

BD176710 1997 bp DNA linear PAT 18-MAR-2003
LOCUS BD176710
DEFINITION Method of examining allergic disease.
ACCESSION BD176710
VERSION BD176710.1 GI:29122420
KEYWORDS WO 02075304-A/11.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1997)
Sugita,Y., Heishi,M., Kagaya,S., Gunji,S. and Tsujimoto,G.
Method of examining allergic disease
Patent: WO 02075304-A 11 26-SEP-2002;
GENOX RESEARCH INC, JAPAN AS REPRESENTED BY GENERAL DIRECTOR OF
NATIONAL CHILDREN'S HOSPITAL, HIDETO YONEKURA, YASUHIRO
YAMAMOTO, SHIGERU SAKURAI, TAKUO WATANABE YUJI SUGITA, MASAYUKI
HEISHI, SHINJI KAGAYA, SHIGEMICHI GUNJI, GOZO TSUJIMOTO
COMMENT OS Homo sapiens (human)
PN WO 02075304-A/11
PD 26-SEP-2002
PR 01-MAR-2002 WO 2002JP001916
PI 21-MAR-2001 JP 01P 081028
PI YUJI SUGITA, MASAYUKI HEISHI, SHINJI KAGAYA, SHIGEMICHI GUNJI, PI
GOZO TSUJIMOTO
PC GO1N33/15, GO1N33/50, C12N15/09, C12Q1/02, C12Q1/68, A61K45/00, PC
A61K31/711,
PC A61K39/395, A61P37/08, A61K48/00, A01M67/027
CC Method of examining allergic disease
FH Key Location/Qualifiers
FT CDS (266) . (1789) .
FEATURES
source 1. 1997
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 372 a 631 c 676 g 318 t
ORIGIN
Query Match 98.2%; Score 1963.4; DB 6; Length 1997;
Best Local Similarity 99.8%; Pred. No. 9,4e-284;
Matches 1987; Conservative 0; Mismatches 1; Indels 3; Gaps 2;
1 CTCGCTCCAGTTGTGACGCGGCGGCTCGGGGTGTGACGCGGCTCGGGAGGCC 60
8 CTCGCTCCAGTTGTGACGCGGCGGCGGCTCGGGGTGTGACGCGGCTCGGGAGGCC 67
61 TCTGTGGGGGGGGGGG--GCGGCTGGGGGGGCGCCCTGTGAGCAGAAACAGAGAA 118
68 TCTGTGGGGGGGGGGGCGGCGGCGGCTCGGGGGGCGCCCTGTGAGCAGAAACAGAGAA 127
119 CGAGCTGGGTGCTGACGAGCAGCAGCTCTTACCTCTGTGTGACGCGGCTGGCTGTGG 178
128 CGAGCTGGGTGCTGACGAGCAGCAGCTCTTACCTCTGTGTGACGCGGCTGGCTGTGG 187
179 CAGGCAATTCAGGAGGTCGCGGAGCTGTGACCACTTGTGACAGTGTGCTCTCACTGCTTC 238
188 CAGGCAATTCAGGAGGTCGCGGAGCTGTGACCACTTGTGACAGTGTGCTCTCACTGCTTC 247
239 AGTTTCCCTGTGGGGGGGATGCGGGGGAGGCTCTGTGTTTCTGTGGCGGCAATTTCA 298
248 AGTTTCCCTGT--GGGGGGGATGCGGGGGAGGCTCTGTGTTTCTGTGGCGGCAATTTCA 306
299 CGGCTGTGATTTCTGTGAGAACTTCCCGGGGTGAGCCCGCTTCTCCAGCTGGCA 358
307 CGGCTGTGATTTCTGTGAGAACTTCCCGGGGTGAGCCCGCTTCTCCAGCTGGCA 366
359 CCGCCCTCCGCTCTCAGCAGGATGCAACGAGGGGCTGGCCCGGGGACCACTGAT 418
367 CCGCCCTCCGCTCTCAGCAGGATGCAACGAGGGGCTGGCCCGGGGACCACTGAT 426
419 CACCAATTCAGACACACCGCCCAAGCAGGGGAGCTGACCTTCCGCAAGGGGACGT 478
|||||

427 CACCAATTCAGACACACCGCCCAAGCAGGGGAGCTGCTTCCGCAAGGGGACGT 486
479 GGTACCACTCTGTGAGGCTGTGAGAAACAGCTGTGACCGGCTGACGACCAACAG 538
|||||
487 GGTACCACTCTGTGAGGCTGTGAGAAACAGCTGTGACCGGCTGACGACCAACAG 546
539 TGGAGAGAGGGGCTGTGAGGCTGTGAGGCGGCTGGGAGGAGGAGGCGCTTCCGCA 598
|||||
547 TGGAGAGAGGGGCTGTGAGGCTGTGAGGCGGCTGGGAGGAGGAGGCGCTTCCGCA 606
599 CCGCAAGCTCAGACCTCATCCGCTGTTCACGGGAGATCTCGGGGACAGAGCTGTCCA 658
607 CCGCAAGCTCAGACCTCATCCGCTGTTCACGGGAGATCTCGGGGACAGAGCTGTCCA 666
659 GCAGTGTGAGCTCCCGAGAGAGTGGGCTTCTGTTGGGAGTCCGGGCGCAACCGCG 718
667 GCAGTGTGAGCTCCCGAGAGAGTGGGCTTCTGTTGGGAGTCCGGGCGCAACCGCG 726
719 GCAGTGTGAGCTCCCGAGAGAGTGGGCTTCTGTTGGGAGTCCAGTACCGGCTGTGAC 778
727 GCAGTGTGAGCTCCCGAGAGAGTGGGCTTCTGTTGGGAGTCCAGTACCGGCTGTGAC 786
779 CGAGGCGCACTCATCATGATGAGGCGCTGTTCCTGTGCAACTCATGATGATGTGGA 838
787 CGAGGCGCACTCATCATGATGAGGCGCTGTTCCTGTGCAACTCATGATGATGTGGA 846
839 GCATTCACGACAGGAGGCGCTGTGACCAAGCTGTGAGAGACCAAGGAGGAGAA 898
847 GCATTCACGACAGGAGGCGCTGTGACCAAGCTGTGAGAGACCAAGGAGGAGAA 906
899 CGGAGCAAGTCCGCGGAGAGAGAGTGGGCGGCGGCTGTGATTAACCTGTGAC 958
907 CGGAGCAAGTCCGCGGAGAGAGAGTGGGCGGCGGCTGTGATTAACCTGTGAC 966
959 TTTGACATTTGGAGACACAGATGCGAGAGAGAGTTTGGAGCTGTCTGACGGTGTG 1018
967 TTTGACATTTGGAGACACAGATGCGAGAGAGAGTTTGGAGCTGTCTGACGGTGTG 1026
1019 CTTGGGGGAAAAAGTGGCGGCTGTGAAATATCAAGTGTGATGACGCGGCTTCT 1078
1027 CTTGGGGGAAAAAGTGGCGGCTGTGAAATATCAAGTGTGATGACGCGGCTTCT 1086
1079 GGAGCAGAGGCGGCTGTGACGAGAGATCAACACGAGAACTGTGCTTCTGAGGCT 1138
1087 GGAGCAGAGGCGGCTGTGACGAGAGATCAACACGAGAACTGTGCTTCTGAGGCT 1146
1139 GATCTGTGACAGGAGGCTGTGATGATGAGAGCAGTGTGACAGAGGAGACCTGTGGA 1198
1147 GATCTGTGACAGGAGGCTGTGATGATGAGAGCAGTGTGACAGAGGAGACCTGTGGA 1206
1199 CTTTCTGGGAGACCGGGGCTGAGCGCTGTGAAACCGCTGTGACAGTGTGCT 1258
1207 CTTTCTGGGAGACCGGGGCTGAGCGCTGTGAAACCGCTGTGACAGTGTGCT 1266
1259 GCAGTGTGCGGAGGAGTGTGATGAGAGAGAGTGTGAGAGCTGTGACAGGAGCTGTG 1318
1267 GCAGTGTGCGGAGGAGTGTGATGAGAGAGAGTGTGAGAGCTGTGACAGGAGCTGTG 1326
1319 CGCGCGCAACATCTGTGTCTGAGAGACCTGTGAGAGAGTGTGAGAGCTGTGAGC 1378
1327 CGCGCGCAACATCTGTGTCTGAGAGACCTGTGAGAGAGTGTGAGAGCTGTGAGC 1386
1379 CAAAGCGAGGAGGAGGCTGTGATGAGAGAGGCTGTGAGAGAGTGTGAGAGGAGGAG 1438
1387 CAAAGCGAGGAGGAGGCTGTGATGAGAGAGGCTGTGAGAGAGTGTGAGAGGAGGAG 1446
1439 GCGCTCTCAACACGAGGAGGAGTGTGAGAGAGGAGTGTGAGAGTGTGAGAGTGTG 1498
1447 GCGCTCTCAACACGAGGAGGAGTGTGAGAGAGGAGTGTGAGAGTGTGAGAGTGTG 1506
1499 CTGGAGAGTCTTCTATATGAGAGGAGTGTGAGAGAGTGTGAGAGTGTGAGAGTGTG 1558
1507 CTGGAGAGTCTTCTATATGAGAGGAGTGTGAGAGAGTGTGAGAGTGTGAGAGTGTG 1566
|||||

QY 1559 GGAAGCCGTGAGAAAGGGGATACCCATGGAACCCCGAGGGCTGTCCAGGCCCCCTCA 1618
 DB 1567 GGAAGCCGTGAGAAAGGGGATACCCATGGAACCCCGAGGGCTGTCCAGGCCCCCTCA 1626
 QY 1619 CGTCTATGAGAGAGCTGTGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1678
 DB 1627 CGTCTATGAGAGAGCTGTGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1686
 QY 1679 GGGCGAGAGAGCTGTGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1728
 DB 1687 GGGCGAGAGAGCTGTGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1746
 QY 1739 GGAAGCCGAGAGCTGTCCAGCTGTGCCCCGAGAGAGAGAGAGAGAGAGAGAGAG 1798
 DB 1747 GGAAGCCGAGAGCTGTCCAGCTGTGCCCCGAGAGAGAGAGAGAGAGAGAGAGAG 1806
 QY 1799 CTTGGCCCCAG 1858
 DB 1807 CTTGGCCCCAG 1866
 QY 1859 GGAAGGATCCAG 1918
 DB 1867 GGAAGGATCCAG 1926
 QY 1919 GGGGCTGTGGGAG 1978
 DB 1927 GGGGCTGTGGGAG 1986
 QY 1979 GGATTCTAAG 1989
 DB 1987 GGATTCTAAG 1997

RESULT 6
 LOCUS AR084686 1987 bp DNA linear PAT 01-SEP-2000
 DEFINITION Sequence 1 from patent US 5981201.
 ACCESSION AR084686
 VERSION AR084686.1 GI:10011456
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1987)
 AUTHORS Avraham, H. and Groopman, J. E.
 TITLE Methods of detection and treatment of breast cancer
 JOURNAL Patent: US 5981201-A 1 09-NOV-1999;
 FEATURES location/Qualifiers
 source 1. 1987
 BASE COUNT 369 a 628 c 672 g 318 t
 ORIGIN

Query Match 95.5%; Score 1909.8; DB 6; Length 1987;
 Best Local Similarity 99.2%; Pred. No. 9.5e-276;
 Matches 1973; Conservative 0; Mismatches 7; Indels 9; Gaps 5;

QY 1 CTCGCTCCAGATTGTGACAGCCGAGACCGCTCGGGGTGTGACAGCCGGCTGCGGAGGCC 60
 DB 8 CTCGCTCCAGATTGTGACAGCCGAGACCGCTCGGGGTGTGACAGCCGGCTGCGGAGGCC 67
 QY 61 TCCCTGG 120
 DB 68 TCCCTGG 127
 QY 121 AGGCTGGGTCCAGTGGACACCACTCCCTACCTCTGTGTCAGCCGGCTGGGCTGTGGCA 180
 DB 128 AGGCTGGGTCCAGTGGACACCACTCCCTACCTCTGTGTCAGCCGGCTGGGCTGTGGCA 187
 QY 181 GGGCAATTCACAGGCTGCGGAGCTGTGACCACTTGTGAGTGTCTCTACCTGCTCAG 240
 DB 188 GGGCAATTCACAGGCTGCGGAGCTGTGACCACTTGTGAGTGTCTCTACCTGCTCAG 247

QY 241 TTTCCTCTGGGGGGCGATGCGGGGAGAGCTCTGTGTTTCTGTGGGCGCATTTACG 300
 DB 248 TTTCCTCTGGGGGGCGATGCGGGGAGAGCTCTGTGTTTCTGTGGGCGCATTTACG 305
 QY 301 GCTGTGATTTCTGTGAGAACTTCCCGGGTGAAGCCCGGTTCTCCAGCTGGGACC 360
 DB 306 GCTGTGATTTCTGTGAGAACTTCCCGGGTGAAGCCCGGTTCTCCAGCTGGGACC 365
 QY 361 CCCCTCCGCTCTGAGAGAGTGCACAGAGGGGCTGGGCCCCGGGAGCCAGTGTATCA 420
 DB 366 CCCCTCCGCTCTGAGAGAGTGCACAGAGGGGCTGGGCCCCGGGAGCCAGTGTATCA 425
 QY 421 CCAATGCGAGACACCCGCCGCCAAGCCAGGGGAGTGGCTTCTCCGAGGGGAGAGCTGG 480
 DB 426 CCAATGCGAGACACCCGCCGCCAAGCCAGGGGAGTGGCTTCTCCGAGGGGAGAGCTGG 485
 QY 481 TCACCATCTGTGAGAGCTGTGAGAAAGAGCTGTACCGGCTCAGCACACACAGTG 540
 DB 486 TCACCATCTGTGAGAGCTGTGAGAAAGAGCTGTACCGGCTCAGCACACACAGTG 545
 QY 541 GACAGAGGGGCTGTGGAGCTGTGGGAGCTGTGGGAGAGGGGAGGCTCTCCGACAGCC 600
 DB 546 GACAGAGGGGCTGTGGAGCTGTGGGAGCTGTGGGAGAGGGGAGGCTCTCCGACAGCC 605
 QY 601 CCAAGCTCAGCTCTATGCTCGGTGTCCAGGGGAAAGATCTCGGGCCAGAGAGCTGTCCAGC 660
 DB 606 CCAAGCTCAGCTCTATGCTCGGTGTCCAGGGGAAAGATCTCGGGCCAGAGAGCTGTCCAGC 665
 QY 661 AGCTGACGCTCCCGAGAGAGGCTGTCTGTGGGAGAGTCCGCGCCACCCGGGG 720
 DB 666 AGCTGACGCTCCCGAGAGAGGCTGTCTGTGGGAGAGTCCGCGCCACCCGGGG 725
 QY 721 ACTACGCTGTGTGAGTGTGGGAGCTGTGGGAGAGCTCATCTACCGGCTGTGACAGCG 780
 DB 726 ACTACGCTGTGTGAGTGTGGGAGCTGTGGGAGAGCTCATCTACCGGCTGTGACAGCG 785
 QY 781 ACGGCCACCTCAGCAATGATGAGAGCGCTGTCTCTGCAACCTCATGAGATGATGAGAGC 840
 DB 786 ACGGCCACCTCAGCAATGATGAGAGCGCTGTCTCTGCAACCTCATGAGATGATGAGAGC 845
 QY 841 ATTACAGCAAGCAAGAGGGGCTATCTGTGACCAAGCTGTGTGAGACCAAGAGCAAGCAAGC 900
 DB 846 ATTACAGCAAGCAAGAGGGGCTATCTGTGACCAAGCTGTGTGAGACCAAGAGCAAGCAAGC 905
 QY 901 GGAACCAAGTGGGCGGAGAGAGAGCTGTGGGAGGGGCTGTACTGTAACCTCAGCAT 960
 DB 906 GGAACCAAGTGGGCGGAGAGAGAGCTGTGGGAGGGGCTGTACTGTAACCTCAGCAT 965
 QY 961 TGAATTTGGGAGACAGAGATGAGAGAGAGATTTGGAGCTGTGTGAGAGGATGAGTACC 1020
 DB 966 TGAATTTGGGAGACAGAGATGAGAGAGAGATTTGGAGCTGTGTGAGAGGATGAGTACC 1025
 QY 1021 TGGGGCAAAAGTGGGCGGCTGTGAATAATATCAAGTGTATGTGAACAGCCAGGCTTCTGTG 1080
 DB 1026 TGGGGCAAAAGTGGGCGGCTGTGAATAATATCAAGTGTATGTGAACAGCCAGGCTTCTGTG 1085
 QY 1081 ACGAGAGGCGGCTGTATGAGAGAGATGACACAGAGAACTGGTGGCTGTCTGGGGGTGA 1140
 DB 1086 ACGAGAGGCGGCTGTATGAGAGAGATGACACAGAGAACTGGTGGCTGTCTGGGGGTGA 1145
 QY 1141 TCCGTGACAGAGGGGCTGTATGATGTATGAGAGAGCTGTGAAGAGGCAACCTGGTAAC 1200
 DB 1146 TCCGTGACAGAGGGGCTGTATGATGTATGAGAGAGCTGTGAAGAGGCAACCTGGTAAC 1205
 QY 1201 TTCTGGGAGAGCGGGGCTGTGAGAGCTGTGTGAACAGCCCTCAGCTCTGTGAGTCTGTG 1260
 DB 1206 TTCTGGGAGAGCGGGGCTGTGAGAGCTGTGTGAACAGCCCTCAGCTCTGTGAGTCTGTG 1265
 QY 1261 ACGTGGCCAGAGGAGATGAGATGAGATGAGAGAGAGCTGTGTGAGAGAGAGAGAGAGAG 1320
 DB 1266 ACGTGGCCAGAGGAGATGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1325

OY	1321	CCGCGAACATCTGTGTCTCAGAGGACCCGTGGGCGCAAGTCAGAGACTTGGCTGTGGCA	1380
Db	1326	CCCGGAACATCTGTGTCTCAGAGGACCCGTGGGCGCAAGTCAGAGACTTGGCTGTGGCA	1385
OY	1381	AAGCGGAGCGGAAGGGGCTAGATCTCAAGCCCGCTGCCCTCAAGTCAGAGCGGCGCCGAG	1440
Db	1386	AAGCGGAGCGGAAGGGGCTAGATCTCAAGCCCGCTGCCCTCAAGTCAGAGCGGCGCCGAG	1445
OY	1441	CTCTGAAAACACGGGCAAGTTCACACAGCAAGTGGATGTCTGCAATTTTGGGCTGCTCT	1500
Db	1446	CTCTGAAAACACGGG---GTTCACACAGCAAGTGGATGTCTGCAATTTTGGGCTGCTCT	1502
OY	1501	GGAGAGCTTCTCATATGAGAGGGGCTCCGTACCCCTAAATGTCATCTGAAAGAGGTGTGG	1560
Db	1503	GGAGAGCTTCTCATATGAGAGGGGCTCCGTACCCCTAAATATTCATCTGAAAGAGGTGTGG	1562
OY	1561	AGGCGCGTGGAGAGAGGGGTACCCGATGGAACCCCGGAGAGGCTGTTCAGAGCCCGTGCACG	1620
Db	1563	AGGCGCGTGGAGAGAGGGGTACCCGATGGAACCCCGGAGAGGCTGTTCAGAGCCCGTGCACG	1622
OY	1621	TCTCTGATGAGAGGTGTGCTGGGAGGACAGAGCCCGCGCGCCACACCTTCGGCAAACTGG	1680
Db	1623	TCTCTGATGAGAGGTGTGCTGGGAGGACAGAGCCCGCGCGCCACACCTTCGGCAAACTGG	1681
OY	1681	CCGAGAAAGCTGGCCCGGGAGCTACGACGATGCAAGTGTGCCACACCTTCGGTCTCAGAGGCAAG	1740
Db	1682	CCGAGAAAGCTGGCCCGGGAGCTACGACGATGCAAGTGTGCCACACCTTCGGTCTCAGAGGCAAG	1741
OY	1741	ACGGCGAGCGGTCTCACCCTCGGCCCGGAAAGCCAGAGAGCCCTGACCCCAACCCGGTGGGGCCTT	1800
Db	1742	ACGGCGAGCGG---TCCACCTCGGCCCGGAAAGCCAGAGAGCCCTGACCCCAACCCGGT---GGCCCTT	1798
OY	1801	TGGCGCCACAGAGACCCAGAGAGTAGTAGTGCAGGCTGCGGCGTGGGGGCACTGACACAGGCCCAAG	1860
Db	1799	TGGCGCCACAGAGACCCAGAGAGTAGTAGTGCAGGCTGCGGCGTGGGGGCACTGACACAGGCCCAAG	1858
OY	1861	AGGGTCCACAGCGGGGCGCAAGTCACTCTCTGTGGTGTGCCCAACAGCAGGGGGCTTGGCCACGTAGGG	1920
Db	1859	AGGGTCCACAGCGGGGCGCAAGTCACTCTCTGTGGTGTGCCCAACAGCAGGGGGCTTGGCCACGTAGGG	1918
OY	1921	GGCTCTGTGGCGCCCGCTGTGACACCCCAAGACTGTGGAAGATGATCGCCCATTAAGACGG	1980
Db	1919	GGCTCTGTGGCGCCCGCTGTGACACCCCAAGACTGTGGAAGATGATCGCCCATTAAGACGG	1978
OY	1981	ATTCTTAAGG 1989	
Db	1979	ATTCTTAAGG 1987	

```
RESULT 7  
AR204704 LOCUS 1987 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 1 from patent US 6368796.  
ACCESSION AR204704  
VERSION AR204704.1 GI:21502099  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
I (bases 1 to 1987)  
AUTHORS Avraham,H. and Groopman,J.E.  
TITLE Methods of detection and treatment of breast cancer  
JOURNAL Patent: US 6368796-A 1 09-APR-2002;  
FEATURES location/qualifiers  
          source l..1987 /organism="unknown"  
BASE COUNT 369 a 628 c 672 g 318 t -  
ORIGIN
```

Query Match 95.5%; Score 1909.8; DB 6; Length 1987;
Best Local Similarity 99.2%; Pred. No. 9.5e+276;
Matches 1973; Conservative 0; Mismatches 7; Indels 9; Gaps 5;

[illegible]

361 CCCCTCCGCTCTACAGCAGATGCCAGAGGCGCTGGGCCCGGGGCAACCCAGTATATCA 420
 366 CCCCTCCGCTCTACAGCAGATGCCAGAGGCGCTGGGCCCGGGGCAACCCAGTATATCA 425
 421 CCAATTCGAGACACCCGCCCAAGCAGGAGAGCTGGCTTCCGCAAGGGGCAAGTGG 480
 426 CCAATTCGAGACACCCGCCCAAGCAGGAGAGCTGGCTTCCGCAAGGGGCAAGTGG 485
 481 TCACATCTCGAGGCGCTGGAGACAGAGCTGGTACCGCTCAAGACACCAACAGAGG 540
 486 TCACATCTCGAGGCGCTGGAGACAGAGCTGGTACCGCTCAAGACACCAACAGAGG 545
 541 GACAGAGAGGCGCTGGAGAGCTGGGAGCGCTCGGAGAGGAGGAGGCGCTTCCGAGAGC 600
 546 GACAGAGAGGCGCTGGAGAGCTGGGAGCGCTCGGAGAGGAGGAGGCGCTTCCGAGAGC 605
 601 CCAAGCTCAGCTCATTCGCGTGGTTCACAGGAGAGATCTCGGCGCAAGAGGCTGTCCAGC 660
 606 CCAAGCTCAGCTCATTCGCGTGGTTCACAGGAGAGATCTCGGCGCAAGAGGCTGTCCAGC 665
 661 AGCTCAGGCTCCGAGAGATGGGCGTTCCTGGTGGGAGGAGTCCGCGCGCCACCCCGGCG 720
 666 AGCTCAGGCTCCGAGAGATGGGCGTTCCTGGTGGGAGGAGTCCGCGCGCCACCCCGGCG 725
 721 ACTAGCTCCTGTGCGTGAAGCTTTGGCCGAGCGTCACTACACCGCGTGTGCACCGCG 780
 726 ACTAGCTCCTGTGCGTGAAGCTTTGGCCGAGCGTCACTACACCGCGTGTGCACCGCG 785
 781 ACGGCGACCTCAATTCGATGAGCGCGTGTCTTGCACCTCATTCGACATGTGTGAGC 840
 786 ACGGCGACCTCAATTCGATGAGCGCGTGTCTTGCACCTCATTCGACATGTGTGAGC 845
 841 ATTACAGCAAGAGACAGGCGCTATCTGACACAGCTGGTGAACCAAGCGGAACACAG 900
 846 ATTACAGCAAGAGACAGGCGCTATCTGACACAGCTGGTGAACCAAGCGGAACACAG 905
 901 GGACCAAGTGGCGCCGAGAGAGAGTGGCGAGGCGGCGTGTACTGAACCTGCAGCATT 960
 906 GGACCAAGTGGCGCCGAGAGAGAGTGGCGAGGCGGCGTGTACTGAACCTGCAGCATT 965
 961 TGACATTGGGAGACACAGATCGAGAGGAGAGTGGAGCTTCTCGACAGGCTGAGTACC 1020
 966 TGACATTGGGAGACACAGATCGAGAGGAGAGTGGAGCTTCTCGACAGGCTGAGTACC 1025
 1021 TGGGCGCAAAAGGTGGCGCGTGAAGAAATATCAAGTGTGATGACGCCAGGCGCTTCCG 1080
 1026 TGGGCGCAAAAGGTGGCGCGTGAAGAAATATCAAGTGTGATGACGCCAGGCGCTTCCG 1085
 1081 ACGAGACGCGCTCATGACGAAGATGCAACAGAGAACTGGTGGCTCTCGTGGGCGTGA 1140
 1086 ACGAGACGCGCTCATGACGAAGATGCAACAGAGAACTGGTGGCTCTCGTGGGCGTGA 1145
 1141 TCCTCACCAGGCGCTGTATCTGTATGTGATGAGACAGTGAAGCAAGCAACCTGTGTACT 1200
 1146 TCCTCACCAGGCGCTGTATCTGTATGTGATGAGACAGTGAAGCAAGCAACCTGTGTACT 1205
 1201 TTCTCGGAGACCGGCGGTGAGAGCTGTGTAACACCGCTAGTCTCTGCAAGTTTCTCTGC 1260
 1206 TTCTCGGAGACCGGCGGTGAGAGCTGTGTAACACCGCTAGTCTCTGCAAGTTTCTCTGC 1265
 1261 ACGTGGCGGAGGAGTGAAGTACTGTGAGAGCAAGAAGCTTGTGACCGGCAAGCTGGCGG 1320
 1266 ACGTGGCGGAGGAGTGAAGTACTGTGAGAGCAAGAAGCTTGTGACCGGCAAGCTGGCGG 1325
 1321 CCGGGAACATCTGTGTCTCAGAGAGCTGTGTGAGAGCTGTGTGAGAGCTTGTGGCTGGCA 1380
 1326 CCGGGAACATCTGTGTCTCAGAGAGCTGTGTGAGAGCTGTGTGAGAGCTTGTGGCTGGCA 1385
 1381 AAGCGAGCGAGAGGAGTGAAGTCAAGCGGCGCTCCGCTCAAGTGTGAGCGCGCGAGG 1440
 1386 AAGCGAGCGAGAGGAGTGAAGTCAAGCGGCGCTCCGCTCAAGTGTGAGCGCGCGAGG 1445
 1441 CTCTCAAAACAGCGGAAGTTCAACAGCAAGTGTGATGTCTGAGAGTTTGGGCTGTCTCT 1500

1446 CTCTCAAAACAGG---GTTACACAGCAAGTGGATGTCTGGAGTTTGGGGGTGTCTCT 1502
 1501 GGGAGGCTTCTCATATATGAGAGGCGCTCCGTACCTCTAAATGTCACTGAAGAAGTGTGG 1560
 1503 GGGAGGCTTCTCATATATGAGAGGCGCTCCGTACCTCTAAATGTCACTGAAGAAGTGTGG 1562
 1561 AGGCGGTGAGAGGAGGAGTACCGCATGGAACCCCGGAGGCGTGTCCAGGCGCGGTGCAGC 1620
 1563 AGGCGGTGAGAGGAGGAGTACCGCATGGAACCCCGGAGGCGTGTCCAGGCGCGGTGCAGC 1622
 1621 TCCTCATGAGCAGCTGTGGAGAGCAGAGCCCGCGCGGCGGCACTTCCGCAAACTGG 1680
 1623 TCCTCATGAGCAGCTGTGGAGAGCAGAGCCCGCGCGGCGGCACTTCCGCAAACTGG 1681
 1681 CCGAGAGCTGGCGCGGAGCTACGACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1740
 1682 CCGAGAGCTGGCGCGGAGCTACGACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1741
 1741 ACGCGGAGGCGTTCACCTTCGCGCGGAGGAGAGCCGCTGACCCGCGGAGGCGGCGCT 1800
 1742 ACGCGGAGGCGTTCACCTTCGCGCGGAGGAGAGCCGCTGACCCGCGGAGGCGGCGCT 1798
 1801 TGGCCCAAGAGAGCAGAGAGTGAAGTGGCGGCGTGGGCGCACTGACAGAGCCCAAG 1860
 1799 TGGCCCAAGAGAGCAGAGAGTGAAGTGGCGGCGTGGGCGCACTGACAGAGCCCAAG 1858
 1861 AGGCTCAGAGCGCGGAGAGTCACTCTCTGTGTGTGAGTGTGAGTGTGAGTGTGAG 1920
 1859 AGGCTCAGAGCGCGGAGAGTCACTCTCTGTGTGTGAGTGTGAGTGTGAGTGTGAG 1918
 1921 GGCCTGGGCGCGCGGTGGAGACAGCCAGACCTGTGCAAGAGATGTGCGCGATTAAGACG 1980
 1919 GGCCTGGGCGCGCGGTGGAGACAGCCAGACCTGTGCAAGAGATGTGCGCGATTAAGACG 1978
 1981 ATTCCTAAG 1989
 1979 ATTCCTAAG 1987

RESULT 9
 HSHYLRK
 LOCUS HSHYLRK 1968 bp mRNA linear PRI 10-OCT-1994
 DEFINITION H. sapiens HYL tyrosine kinase mRNA.
 ACCESSION X77278
 VERSION X77278.1 GI:471312
 KEYWORDS HYLTK gene; nonreceptor protein tyrosine kinase.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 Sakano, S., Iwama, A., Inazawa, J., Ariyama, T., Ohno, M. and Suda, T.
 Molecular cloning of a novel non-receptor tyrosine kinase, HYL
 (hematopoietic consensus tyrosine-lacking kinase)
 Oncogene 9 (4), 1155-1161 (1994)

JOURNAL MEDLINE 94181267
 PUBMED 8134117
 REFERENCE 2 (bases 1 to 1968)
 AUTHORS Iwama, A.
 DIRECT SUBMISSION
 TITLE Submitted (14-JAN-1994) A. Iwama, Dept of Cell Differentiation,
 Inst of Mol Embryology & Genetics, Kumamoto University School of
 Medicine, 2-2-1 Honjo, Kumamoto 860, JAPAN

FEATURES
 source location/Qualifiers
 1..1968
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="19p13"
 /tissue_type="hematopoietic cell"
 /clone_lib="Ur-7 cDNA library"
 1..1968

gene

mrna
CDS
/gene="HHYLTk"
1..1968
/gene="HHYLTk"
208..1731
/gene="HHYLTk"
/codon.start=1
/product="Hyl tyrosine kinase"
/protein_id="CAA54493.1"
/db_xref="GI:557272"
/translacion="MAGRSLSVSRFAHGDSDABELPRSPFLRAHPPVARSAMPT
RRAPGTCICHTKCHTRPKRDELAFKRGDVTLLIACENKSHRYVKKHTSGESILLAL
GALPERALSADPKLSMPFHGKISQEAVALQIDPEDELFLVRESARHGDVILV
SFRGDIYHRLHNDGHLTIDEAFECNLMDWEHYSKIDGALCTKLRKRKHGTS
AEELARAGMLNLQHLTLAGQIDEGEVALQIDPEDELFLVRESARHGDVILV
TAVMTKQHEMLYLLGLVILHOGLYIVMEHYSKIDGALCTKLRKRKHGTS
HVAEMKLEYSKILHRLDIAANILVEDILVAKSDGLKAEKRGDLSRRLPYKMTA
PEALKHGFSTKSDVMSFSGVILMEVSTYGRAPYPMKSLKEVSEAVEMGEPEP
GPVHVLMSCHMEPEPARPPFRKIAELKELRLSAGAPASVSGDDADGSTSPNSQEP"
BASE COUNT 391 a 614 c 652 g 311 t
ORIGIN
Query Match 95.5% Score 1909.4; DB 9; Length 1968;
Best Local Similarity 99.7% Pred. No. 1.le-275;
Matches 1945; Conservative 0; Mismatches 1; Indels 5; Gaps 3;
52 CGGAGGCTCTCTGCGGGGCGCGGGG--GCGGCTCGGGGGCCGCCCTGAGCAGAAA 109
1 CGGAGGCTCTCTGCGGGGCGCGGGGCGCGGCTCGGGGGCGCCCTGAGCAGAAA 60
110 CAGGAACACGAGCTGGGTCCAGTGGCAGCCAGCTCTCTCTCTCTCTCTCTCTCTCT 169
61 CAGGAACACGAGCTGGGTCCAGTGGCAGCCAGCTCTCTCTCTCTCTCTCTCTCTCT 120
170 GGCCTGGGAGGCTATCCAGGCTCCCGAGCTGTAGCAGCTGTAGTGTAGTGTAGTGT 229
121 GGCCTGGGAGGCTATCCAGGCTCCCGAGCTGTAGCAGCTGTAGTGTAGTGTAGTGT 180
QY 230 ACCTGCTCAGTTTCTCTGCGGGGCGAGTGGCGGGGCGAGGCTCTCTGTTTCTGCG 289
Db 181 ACCTGCTCAGTTTCTCTGCGGGGCGAGTGGCGGGGCGAGGCTCTCTGTTTCTGCG 239
QY 290 GGCATTTCAGGCTGATCTCTGTAGAACTTCCCGGGGTGAGCCCGGCTTCTCTCG 349
Db 240 GGCATTTCAGGCTGATCTCTGTAGAACTTCCCGGGGTGAGCCCGGCTTCTCTCG 299
QY 350 AGCTGGACCCCTCTCCGTTCTAGCAGAGATGCCAAGAGGGCTGGGCGCGGCGAC 409
Db 300 AGCTGGACCCCTCTCCGTTCTAGCAGAGATGCCAAGAGGGCTGGGCGCGGCGAC 359
QY 410 CCAAGTATACCAATGCAAGCACACCCGCCCAAGCAGGGAGCTGCGCTTCCGAA 469
Db 360 CCAAGTATACCAATGCAAGCACACCCGCCCAAGCAGGGAGCTGCGCTTCCGAA 419
QY 470 GGGGAGCTGGTACCATCTCTGAGGCTGCGAAGAACAGCTGTACCGGCTCAAGCA 529
Db 420 GGGGAGCTGGTACCATCTCTGAGGCTGCGAAGAACAGCTGTACCGGCTCAAGCA 479
QY 530 CCAACACAGTGAAGAGAGGGCTGTGAGCTGGGCGCTGCGGAGAGGGAGGCGCT 589
Db 480 CCAACACAGTGAAGAGAGGGCTGTGAGCTGGGCGCTGCGGAGAGGGAGGCGCT 539
QY 590 CTCGCGAGACCCCAAGCTCATGCTCATGCTGTTCACGGGAAGATCTCGGCGCAGA 649
Db 540 CTCGCGAGACCCCAAGCTCATGCTCATGCTGTTCACGGGAAGATCTCGGCGCAGA 599
QY 650 GGGTTCACAGCTGAGCTCCGAGAGATGGGCTTCTCTGTGCGGAGTCCGCGCG 709
Db 600 GGGTTCACAGCTGAGCTCCGAGAGATGGGCTTCTCTGTGCGGAGTCCGCGCG 659
QY 710 CCAACCCGAGCTAGCTCTGTGCTGAGACTTTGGCGGAGCTATCAGTACCGCGT 769
Db 660 CCAACCCGAGCTAGCTCTGTGCTGAGACTTTGGCGGAGCTATCAGTACCGCGT 719

QY 770 GCTGCACCGGAGCGCCACCTCACAATGATGAGCCGTCTTCTGCAACCTCATGA 829
Db 720 GCTGCACCGGAGCGCCACCTCACAATGATGAGCCGTCTTCTGCAACCTCATGA 779
QY 830 CATGCTGAGCATTACGACGAAGCAAGGCGCTATCTGCACCAAGCTGTGAGCCAA 889
Db 780 CATGCTGAGCATTACGACGAAGCAAGGCGCTATCTGCACCAAGCTGTGAGCCAA 839
QY 890 GCGGAACACGAGGACCAAGTCCGCGGAGGAGGAGCTGCCAGGCGCGGTGATCAGA 949
Db 840 GCGGAACACGAGGACCAAGTCCGCGGAGGAGGAGCTGCCAGGCGCGGTGATCAGA 899
QY 950 CTTGACACATTGACATTGGAGACACAGATCGAGAGGAGAGTTTGGAGCTGCTGCA 1009
Db 900 CTTGACACATTGACATTGGAGACACAGATCGAGAGGAGAGTTTGGAGCTGCTGCA 959
QY 1010 GGGTGAATACCTGGGCGCAAAAGTGGCCGCTGAGAGATATCAAGTGTATGTACAGCCA 1069
Db 960 GGGTGAATACCTGGGCGCAAAAGTGGCCGCTGAGAGATATCAAGTGTATGTACAGCCA 1019
QY 1070 GGCCTTCTGAGAGAGAGCGCGCTCATGAGAGATGACACAGAGAACTGGTGGCT 1129
Db 1020 GGCCTTCTGAGAGAGAGCGCGCTCATGAGAGATGACACAGAGAACTGGTGGCT 1079
QY 1130 CTTGGGCGTATCTGACACAGGCGCTGTACATTGTATGAGACAGTGTAGCAGGCA 1189
Db 1080 CTTGGGCGTATCTGACACAGGCGCTGTACATTGTATGAGACAGTGTAGCAGGCA 1139
QY 1190 CTTGCTGATCTTCTGCGGACCGGGGCTGACGCCCTGTGTAACACGCTGACTCTGCA 1249
Db 1140 CTTGCTGATCTTCTGCGGACCGGGGCTGACGCCCTGTGTAACACGCTGACTCTGCA 1199
QY 1250 GTTTCCTGACAGTGGCCGAGGCGATGAGTACCTGGAGAGCAAGAGCTTGTACACG 1309
Db 1200 GTTTCCTGACAGTGGCCGAGGCGATGAGTACCTGGAGAGCAAGAGCTTGTACACG 1259
QY 1310 CGACTGCGCGCCGCAACATCTGTGCTCAGAGAGACCTGTGGCCAAAGTCAAGCACTT 1369
Db 1260 CGACTGCGCGCCGCAACATCTGTGCTCAGAGAGACCTGTGGCCAAAGTCAAGCACTT 1319
QY 1370 TGGCTGCGCAAGCCGAGGAGGAGGAGGCTGAGACTCAAGCGGCTGCGCTCAAGTGGAC 1429
Db 1320 TGGCTGCGCAAGCCGAGGAGGAGGAGGCTGAGACTCAAGCGGCTGCGCTCAAGTGGAC 1379
QY 1430 GCGCGCCGAGGCTCTCAACACGGAAGTTCACACAGCAAGTCCGATGTCTGAGTCTTGG 1489
Db 1380 GCGCGCCGAGGCTCTCAACACGGAAGTTCACACAGCAAGTCCGATGTCTGAGTCTTGG 1439
QY 1490 GGTGCTGCTGAGAGGCTTCTCATATGAGAGGCGCTCGTACCTTAATATGTACTGAA 1549
Db 1440 GGTGCTGCTGAGAGGCTTCTCATATGAGAGGCGCTCGTACCTTAATATGTACTGAA 1499
QY 1550 AGAGGTGTGAGAGGCGGTGAGAGAGGCTTCCGATATGAAACCCCGGAGGCGTGTACAG 1609
Db 1500 AGAGGTGTGAGAGGCGGTGAGAGAGGCTTCCGATATGAAACCCCGGAGGCGTGTACAG 1559
QY 1610 CCGCTGACAGTCTCATGAGACAGTGTGAGAGAGAGCCCGCGCGCGCGCACCTT 1669
Db 1560 CCGCTGACAGTCTCATGAGACAGTGTGAGAGAGAGCCCGCGCGCGCGCACCTT 1619
QY 1670 CCGCAACCTGCGGAGAGAGTGGCCCGGAGAGTACCAAGTGCAGTGTGCGCCAGCTCCGT 1729
Db 1620 CCGCAACCTGCGGAGAGAGTGGCCCGGAGAGTACCAAGTGCAGTGTGCGCCAGCTCCGT 1679
QY 1730 CTTAGGCGAGAGAGCGGCGAGCTGCAACCTGCCCAACAGAGAGCTGTGACCCACCC 1789
Db 1680 CTTAGGCGAGAGAGCGGCGAGCTGCAACCTGCCCAACAGAGAGCTGTGACCCACCC 1739
QY 1790 GGTGGGCGCTTGGCCCGAGAGAGAGAGAGTGTGAGAGTGGGCGGCTGAGGAGTGTGAC 1849
Db 1740 GGT--GGCCCTTGGCCCGAGAGAGAGAGAGTGTGAGAGTGGGCGGCTGAGGAGTGTGAC 1797


```

OY 1010 GGGTGAGTACCTGGGGCAAAAGTGGCCGTGAGAAATATCAAGTGTGATGTGACAGCCCA 1069
    |||||||
DB 960 GGGTGAATACCTGGGGGCAAAAGTGGCCGTGAGAAATATCAAGTGTGATGTGACAGCCCA 1019
OY 1070 GGGCTTCTGAGACGAGAGCGCCGTGATGACGAAATGACAGACGAAACCTGTGCGCTCT 1129
    |||||||
DB 1020 GGGCTTCTGAGACGAGAGCGCCGTGATGACGAAATGACAGACGAAACCTGTGCGCTCT 1079
OY 1130 CCTGGGCGTGTATCTTGCACACGAGGGCTTACTATTGTCTATGACACGCTGAGCAGGGCAA 1189
    |||||||
DB 1080 CTTGGGCGTGTATCTTGCACACGAGGGCTTACTATTGTCTATGAGACGCTGAGCAGGGCAA 1139
OY 1190 CTTGGTGAACCTTTTGGGACCGGGGGTGCAGCCCTGCTGAACACCGCTGACGCTCTGCA 1249
    |||||||
DB 1140 CTTGGTGAACCTTTTGGGACCGGGGGTGCAGCCCTGCTGAACACCGCTGACGCTCTGCA 1199
OY 1250 GTTTCCTGTCACGATGGCCGAGGGGATGAGTACTGTGAGAGACAGAAAGCTTGTGCACCG 1309
    |||||||
DB 1200 GTTTCCTGTCACGATGGCCGAGGGGATGAGTACTGTGAGAGACAGAAAGCTTGTGCACCG 1259
OY 1310 CGACTGGCCCGCCCAACATCTCTGTCTCAGAGACCTGTTGGCCAAAGTTCAGGACTT 1369
    |||||||
DB 1260 CGACTGGCCCGCCCAACATCTCTGTCTCAGAGACCTGTTGGCCAAAGTTCAGGACTT 1319
OY 1370 TGGCGTGGCAAAAGCGGAGCGGAGGGGCTGAGTCAAGCGCGGCTGCGCTCAAGTGGAC 1429
    |||||||
DB 1320 TGGCGTGGCAAAAGCGGAGCGGAGGGGCTGAGTCAAGCGCGGCTGCGCTCAAGTGGAC 1379
OY 1430 GGGCGCCGAGGCTCTCAAAACACGAGGAGTTCACGACGAACTGAGTGTCTGAGTCTTGG 1489
    |||||||
DB 1380 GGGCGCCGAGGCTCTCAAAACACGAGGAGTTCACGACGAACTGAGTGTCTGAGTCTTGG 1439
OY 1490 GGTGCTGCTGTGGAGAGTCTTCTCATATGACGGGCTCGTACCTAAATGTCACTGAA 1549
    |||||||
DB 1440 GGTGCTGCTGTGGAGAGTCTTCTCATATGAGAGGGCTCGTACCTAAATGTCACTGAA 1499
OY 1550 AGAGTGTGCGAGGGCGTGGAGAGGGGATGACGATGGAACCCCGGAGGGGCTGTCCAGG 1609
    |||||||
DB 1500 AGAGTGTGCGAGGGCGTGGAGAGGGGATGACGATGGAACCCCGGAGGGGCTGTCCAGG 1559
OY 1610 CCCCGTGCACGCTCTCATATGACGAGCTGTGGAGGACAGAGCCCGCCGCGCACCTT 1669
    |||||||
DB 1560 CCCCGTGCACGCTCTCATATGACGAGCTGTGGAGGACAGAGCCCGCCGCGCACCTT 1619
OY 1670 CCGCAAACTGGCCGAGAAAGCTTGGCCGGAGCTACGACAGTGCAGGTGCCCAAGCTCCGT 1729
    |||||||
DB 1620 CCGCAAACTGGCCGAGAAAGCTTGGCCGGAGCTACGACAGTGCAGGTGCCCAAGCTCCGT 1679
OY 1730 CTCAGGGGACGAGCGCGAGCGGCTGCACCTCGCCCGGAGGCCAGAGCCCTGACCCACCC 1789
    |||||||
DB 1680 CTCAGGGGACGAGCGCGAGCGGCTGCACCTCGCCCGGAGGCCAGAGCCCTGACCCACCC 1739
OY 1790 GGTGGGGCCCTTGGCCCGAGAGGACCGAGAGTGGAGTGGCGCTGGGGGCGCACTGAC 1849
    |||||||
DB 1740 GGT--GGCCCTTGGCCCGAGAGGACCGAGAGTGGAGTGGCGCTGGGGGCGCACTGAC 1797
OY 1850 CAGGCCCAAGAGGGTCCAGGCGGCGCAAGTCACTCTGCTGCTGCCACAGCAGGGGCTGG 1909
    |||||||
DB 1798 CAGGCCCAAGAGGGTCCAGGCGGCGCAAGTCACTCTGCTGCTGCCACAGCAGGGGCTGG 1857
OY 1910 CCCAGTGGGGGGCTCTGGGGGGCGCGTGGAGACCCCGACAGCTTGGGAAGGATGATGCGCC 1969
    |||||||
DB 1858 CCCAGTGGGGGGCTCTGGGGGGCGCGTGGAGACCCCGACAGCTTGGGAAGGATGATGCGCC 1917
OY 1970 GATTAAGAGCGATTTCTAAGGACTCT 1994
    |||||||
DB 1918 GATTAAGAGCGATTTCTAAGGACTCT 1942

```

RESULT 12
HSM802388
LOCUS

HSM802388

1804 bp

mRNA

linear

Pri 18-FEB-2000

```

DEFINITION Homo sapiens mRNA; cDNA DKFZp434N1212 (from clone DKFZp434N1212);
partial cds.
ACCESSION AL137754
VERSION AL137754.1 GI:6808456
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Duesterhoeft,A., Lauber,J., Mewes,H.W., Well,B. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (15-JAN-2000) MIPS, Am Klopfersplitz 18a, D-82152
Martinsried, GERMANY
COMMENT
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp434N1212) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://www.mips.biochem.mpg.de/proj/cDNA/.
Location/Qualifiers
1..1804
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="19p13"
/clone="DKFZp434N1212"
/tissue_type="testis"
/clone_lib="434 (synonym: htes3). Vector pSPORT1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..1804
/gene="DKFZp434N1212"
<1..1574
/protein_id="CAB70906.1"
/note="protein-tyrosine kinase (EC 2.7.1.112) (Homo
sapiens)"
/codon_start=3
/product="hypotheical protein"
/protein_id="CAB70906.1"
/db_xref="GI:6808457"
/translation="GNGATRSKLSGRDLGCSRLAEILLGAGRGSGAPPEVOYSPR
FLRAWHPVSAARMPTRMARVQCTICTCEHTRPPGSLARKGDVITLACEKSW
YRVKHTSGOGLLAAGLRARELSADPKLSLMPFEGKLSGDAVOQLPDEGLF
LVRESARHPDGVLYCVSPGRVIVHYRLVRDGHLLIDAEVFCNLMDVHYSKDGA
ICTKLVRPKRHGTSAPDEELARAGMLNLNHLTGADIGSEFCNLMDVHYSKDGA
VKNIKCDVTAQAFIDETA.VMTKMOHEIN.VRLIGVTLHGGLIYVMEHVSAGNLVNFLE
RGALVNTAQLDLSLHVAEGMEYLESKLVLRDLAARNILVSEDLVAKVDFGLAKA
ERKGLDSSRLPVKWTAPALHKGFTS.SDVMSFVLLMEVFTGRVAPYKMSLEVS
DAVEGGRMEPEEGCPGPVHVLMSCWEAEAPARPPFRKLAEKLAELRSAGAPASVS
GODADGSTSPPRSQEP"
1782
polyA_site
/gene="DKFZp434N1212"
BASE COUNT 364 a 557 c 614 g 269 t
ORIGIN
Query Match 83.5%; Score 1669.2; DB 9; Length 1804;
Best Local Similarity 98.4%; Pred. No. 8.3e-240;
Matches 1686; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
OY 287 GCGGCATTTTCAGGCGCTGTGATTTGCTGAGAGACTTCCCGGGTGAAGCCCGCTTCTCT 346
    |||||
DB 80 GGGGGCGGGCGCGCGCGCTCGGAGCGCCCGCTAGAGAGTGGAGCCCGCTTCTCT 139
    |||||
OY 347 CCGAGCTGGGACCGCCCGCTCCGCTGACGACGAGTGGCAACGAGGCGGTGGGCGCCGGG 406
    |||||
DB 140 CCGAGCTGGGACCGCCCGCTCCGCTGACGACGAGTGGCAACGAGGCGGTGGGCGCCGGG 199
    |||||
OY 407 CACGAGTGTATCACCAATGCGAGACACCGCCCGCAAGCCAGGAGGAGCTGGCTTCG 466
    |||||

```

Db 200 CACCAAGTATACCAAAATGCGAGCACACCCGCCCAAGCCAGGGAGAGTGGCTTCCG 259
Oy 467 CAAGGGGACGTGTGTACCATCTCTGAGAGCCTGCGAGAAACAAGAGCTGTACCCGCTCAA 526
Db 260 CAAGGGGACGTGTGTACCATCTCTGAGAGCCTGCGAGAAACAAGAGCTGTACCCGCTCAA 319
Oy 527 GCACCAACACAGTGTGAGAGAGGGGCTGTGAGAGCTGTGGGCGCTGCGGGAGCCGGAGGC 586
Db 320 GCACCAACACAGTGTGAGAGAGGGGCTGTGAGAGCTGTGGGCGCTGCGGGAGCCGGAGGC 379
Oy 587 CCTCTCGCAGACCCCAAGCTCAGCTTCATCCGTGTGTCCACGGGAAAGATCTCGGGCA 646
Db 380 CCTCTCGCAGACCCCAAGCTCAGCTTCATCCGTGTGTCCACGGGAAAGATCTCGGGCA 439
Oy 647 GAGAGCTGTCCAGAGAGCTGACAGCTCCCGAGAGTGGCTGTCTCTGTCGGGAGTCCGC 706
Db 440 GAGAGCTGTCCAGAGAGCTGACAGCTCCCGAGAGTGGCTGTCTCTGTCGGGAGTCCGC 499
Oy 707 GCGCACCCCGGAGCTACGTCTGTGTGTGAGCTTTGGCGGAGCGTCACTCACTACCG 766
Db 500 GCGCACCCCGGAGCTACGTCTGTGTGTGAGCTTTGGCGGAGCGTCACTCACTACCG 559
Oy 767 CGTGTCTCAGCCGCGAGCGCCACCTCAGCATCTGATGAGGCGCTGTCTCTGTCAGACCTCAT 826
Db 560 CGTGTCTCAGCCGCGAGCGCCACCTCAGCATCTGATGAGGCGCTGTCTCTGTCAGACCTCAT 619
Oy 827 GAGCATGTGTGAGCATTTACAGCAAGAGACAGGCGCTATCTGACCAACAGCTGTGTAGACC 886
Db 620 GAGCATGTGTGAGCATTTACAGCAAGAGACAGGCGCTATCTGACCAACAGCTGTGTAGACC 679
Oy 887 AAACCGGAAACACGAGGAGCAAGTGGCGGAGAGAGAGTGGCGGCGGCTGTGTACT 946
Db 680 AAACCGGAAACACGAGGAGCAAGTGGCGGAGAGAGAGTGGCGGCGGCTGTGTACT 739
Oy 947 GAACCTGACAGATTGACATTGGGAGACAGATGAGAGAGAGATTGGAGCTGTCT 1006
Db 740 GAACCTGACAGATTGACATTGGGAGACAGATGAGAGAGAGATTGGAGCTGTCT 799
Oy 1007 GAGAGGTGATGACTGGGGGAAAAAGTGGCGGCTGTGAACAAATATCAAGTGTGTGAGACAGC 1066
Db 800 GAGAGGTGATGACTGGGGGAAAAAGTGGCGGCTGTGAACAAATATCAAGTGTGTGAGACAGC 859
Oy 1067 CGAGGCTTCTCTGAGACAGAGCGGCTGTATGACGAAAGATCAACACAGAACTGGTGGC 1126
Db 860 CGAGGCTTCTCTGAGACAGAGCGGCTGTATGACGAAAGATCAACACAGAACTGGTGGC 919
Oy 1127 TCTCTCTGAGGCTGTATCTGACCAAGGGGCTGTATGATGAGACAGCTGAGCAAGG 1186
Db 920 TCTCTCTGAGGCTGTATCTGACCAAGGGGCTGTATGATGAGACAGCTGAGCAAGG 979
Oy 1187 CAACCTGTTGAACTTTTGTGGGAGCCGGGGTCCAGCCCTGTGAACACCCCTCAAGTCTCT 1246
Db 980 CAACCTGTTGAACTTTTGTGGGAGCCGGGGTCCAGCCCTGTGAACACCCCTCAAGTCTCT 1039
Oy 1247 GCAATTTTCTCTGACAGTGGCGAGGAGCATGAGTACTGGAGAGCAAGAGCTGTGCA 1306
Db 1040 GCAATTTTCTCTGACAGTGGCGAGGAGCATGAGTACTGGAGAGCAAGAGCTGTGCA 1099
Oy 1307 CCGGAGACTTGGCGCCCGCAACATCTGTGTCTCAGAGAGACCTGTGGCCCAAGTCAAGCA 1366
Db 1100 CCGGAGACTTGGCGCCCGCAACATCTGTGTCTCAGAGAGACCTGTGGCCCAAGTCAAGCA 1159
Oy 1367 CTTTGGGCTTGGCAAAAGCCGAGGAGGAGGCTGTAGACTCAAGCCGCTGCCCTCAAGTG 1426
Db 1160 CTTTGGGCTTGGCAAAAGCCGAGGAGGAGGCTGTAGACTCAAGCCGCTGCCCTCAAGTG 1219
Oy 1427 GAGGCGCCCGGAGGCTCTCAAAACAGGGAGTTTACACAGCAAGTGGAGTCTGTGAGTTT 1486
Db 1220 GAGGCGCCCGGAGGCTCTCAAAACAGGGAGTTTACACAGCAAGTGGAGTCTGTGAGTTT 1279
Oy 1487 TGGGGTGTCTGTGGAGGTTCTTCTCATATGAGAGGGGCTTACCTTAAATGTCTACT 1546
Db 1280 TGGGGTGTCTGTGGAGGTTCTTCTCATATGAGAGGGGCTTACCTTAAATGTCTACT 1339

Oy 1547 GAAAGAGTGTCTGAGAGGCGCTGTGAGAAAGGGGTACCGCATGTGAAACCCCGAGGCTGTCC 1606
Db 1340 GAAAGAGTGTCTGAGAGGCGCTGTGAGAAAGGGGTACCGCATGTGAAACCCCGAGGCTGTCC 1399
Oy 1607 AGGCGCCGCTGTGACGCTCTCATATGACAGCTGTGTGGAGGAGAGAGCCCGCCCGGCAACC 1666
Db 1400 AGGCGCCGCTGTGACGCTCTCATATGACAGCTGTGTGGAGGAGAGAGCCCGCCCGGCAACC 1459
Oy 1667 CTTCCGCAAACTGTGCGCGAGAGTGTGGCCCGGAGCTGACGAGTGTGAGTGTGCCAGGCTC 1726
Db 1460 CTTCCGCAAACTGTGCGCGAGAGTGTGGCCCGGAGCTGACGAGTGTGAGTGTGCCAGGCTC 1519
Oy 1727 CGTCTCAGGCGAGAGCGCGGCTCCACCTGCGCCCGAAGCCAGAGAGCCCTGACCCCA 1786
Db 1520 CGTCTCAGGCGAGAGCGCGGCTCCACCTGCGCCCGAAGCCAGAGAGCCCTGACCCCA 1579
Oy 1787 CCGGCTGGGCGCTTGTGCGCCCGAGAGAGCCGAGAGAGTGTGAGTGTGGGCGGCACT 1846
Db 1580 CCGGCTGGGCGCTTGTGCGCCCGAGAGAGCCGAGAGAGTGTGAGTGTGGGCGGCACT 1639
Oy 1847 GACCAGGCGCCAGAGAGGAGGTCAGAGGCGGCAAGTATCTCTGTGTGCCACAGAGGAGG 1906
Db 1640 GACCAGGCGCCAGAGAGGAGGTCAGAGGCGGCAAGTATCTCTGTGTGCCACAGAGGAGG 1699
Oy 1907 TGGCCACAGTAAAGGAGGCTGTGGGCGGCGGCTGTGACACCCAGAGCTTGAAGATGATG 1966
Db 1700 TGGCCACAGTAAAGGAGGCTGTGGGCGGCGGCTGTGACACCCAGAGCTTGAAGATGATG 1759
Oy 1967 CCGGATAAAGAGGATTTAAAGACTTAAAGAA 2000
Db 1760 CCGGATAAAGAGGATTTAAAGACTTAAAGAA 1793

RESULT 13
AK055395
LOCUS
DEFINITION
Homo sapiens cDNA FLJ30833 fis, clone FEBR200207, highly similar
to MEGAKARYOCYTE-ASSOCIATED TYROSINE-PROTEIN KINASE (EC 2.7.1.112).
AK055395
VERSION
AK055395.1 GI:16550112
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
Kakamori, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hiro, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuna, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
Oshima, A., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K.
and Isogai, T.
NEDO human cDNA sequencing project

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Katsarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan: cDNA full insert sequencing:
Research Association for Biotechnology (RAB): cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.): 5' & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.

COMMENT

FEATURES
source
1. 2073
location/Qualifiers
/organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FEBRA200207"
/tissue_type="brain"
/clone_lib="FEBRA2"
/dev_stage="fetus"
/notes="cloning vector: PME185FL3"

BASE COUNT      415 a      634 c      703 g      321 t
ORIGIN

Query Match      83.2%; Score 1663; DB 9; Length 2073;
Best Local Similarity 99.4%; Pred. No. 6.7e-239;
Matches 1669; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 311 TCGTGAAGAACTTCCCGGGGTGAGCCCGCGCTTCTCCGAGCTGGACCCCGCTCCGCT 370
   |||||
Db 395 TGGTCAAGCAGCAAGCTTCTGTGAGCCCGCGCTTCTCCGAGCTGGACCCCGCTCCGCT 454

QY 371 CTCAGCCAGATGCGCAACGAGCGCTGGGCCCCGGGACCCAGTGATCACCAGATGCGA 430
   |||||
Db 455 CTCAGCCAGATGCGCAACGAGCGCTGGGCCCCGGGACCCAGTGATCACCAGATGCGA 514

QY 431 GCACACCCCGCCCAAGCCAGGAGAGCTGGCTTCCGCAAGGCGAGCTGTACCATCTCT 490
   |||||
Db 515 GCACACCCCGCCCAAGCCAGGAGAGCTGGCTTCCGCAAGGCGAGCTGTACCATCTCT 574

QY 491 GGAGGCGTGCAGAACAGAGCGTGTACCGCGTCAAGCAGCAACACAGTGGACAGAGGG 550
   |||||
Db 575 GGAGGCGTGCAGAACAGAGCGTGTGTACCGCGTCAAGCAGCAACACAGTGGACAGAGGG 634

QY 551 GCTGCTGAGCAGTGGGCGCTGCGGAGCGGAGGCGCTCTCCGAGACCCCAAGCTCAG 610
   |||||
Db 635 GCTGCTGAGCAGTGGGCGCTGCGGAGCGGAGGCGCTCTCCGAGACCCCAAGCTCAG 694

QY 611 CCTCATGCGCGTGTCTCAGCGGAGATCTCGGGCCAGAGGCTGTCCAGCAGCTGACGC 670
   |||||
Db 695 CCTCATGCGCGTGTCTCAGCGGAGATCTCGGGCCAGAGGCTGTCCAGCAGCTGACGC 754

QY 671 TCCCGAGATGGAGCTTCTGCTGCTGGGAGTCCCGGCGCAACCCCGGGGAGTACGTCT 730
   |||||
Db 755 TCCCGAGATGGAGCTTCTGCTGCTGGGAGTCCCGGCGCAACCCCGGGGAGTACGTCT 814

QY 731 GTGCGTGAAGCTTTGGCGCGAGCTATCCACTACCGGCTGTGCACCGCGAGCGCCACT 790
   |||||
Db 815 GTGCGTGAAGCTTTGGCGCGAGCTATCCACTACCGGCTGTGCACCGCGAGCGCCACT 874

QY 791 CACATCGATGAGCGCGTGTCTCTGCAACCTCATGAGCATGTGTGAGCATTTACGCAA 850
   |||||
Db 875 CACATCGATGAGCGCGTGTCTCTGCAACCTCATGAGCATGTGTGAGCATTTACGCAA 934

QY 851 GGACAAGGCGCTATGTGCAACAGCTGTGAGAACCAAGCGGAAACACGGGACCAAGTC 910
   |||||
Db 935 GGACAAGGCGCTATGTGCAACAGCTGTGAGAACCAAGCGGAAACACGGGACCAAGTC 994

QY 911 GGCCGAGAGAGAGCTGGCCAGGGCGGGCTGTACTGAACTGACAGCATTTGACATTTGG 970
   |||||
Db 995 GGCCGAGAGAGAGCTGGCCAGGGCGGGCTGTACTGAACTGACAGCATTTGACATTTGG 1054

QY 971 AACACAGATCGAGAGGAGAGATTTGAGACTGTCTCGAGGGGTGAGTACTTGGGCAAAA 1030
   |||||
Db 1055 AACACAGATCGAGAGGAGAGATTTGAGACTGTCTCGAGGGGTGAGTACTTGGGCAAAA 1114

QY 1031 GGTGGCCGTGAGAAATATCAAGTGTGATGTGACAGGCCAGGGCTTCTGGAGAGAGCGG 1090
   |||||
Db 1115 GGTGGCCGTGAGAAATATCAAGTGTGATGTGACAGGCCAGGGCTTCTGGAGAGAGCGG 1174

QY 1091 CGTCATGAGAGAGATGCAACACAGAACTGTGCGCTCTCTGGGCGGTGATCCTGACCA 1150
   |||||
Db 1175 CGTCATGAGAGAGATGCAACACAGAACTGTGCGCTCTCTGGGCGGTGATCCTGACCA 1234

QY 1151 GGGGCTGTACATTTGTATGAGACAGTGTGAGCAAGGGCAACCTGTGTACTTCTGCGGAC 1210
   |||||
Db 1235 GGGGCTGTACATTTGTATGAGACAGTGTGAGCAAGGGCAACCTGTGTACTTCTGCGGAC 1294

```

```

QY 1211 CCGGGGTGAGCCCTGTGTAACACGCTCAGCTCCTGAGTTTCTCTGACGTGGCGA 1270
   |||||
Db 1295 CCGGGGTGAGCCCTGTGTAACACGCTCAGCTCCTGAGTTTCTCTGACGTGGCGA 1354

QY 1271 GGGCATGAGTACCTGGAGAGCAAGAGCTTGTGACCCGCGAGCTGGCCCGCAACAT 1330
   |||||
Db 1355 GGGCATGAGTACCTGGAGAGCAAGAGCTTGTGACCCGCGAGCTGGCCCGCAACAT 1414

QY 1331 CCTGCTCAGAGAGACCTGTGGCCAGGTCAAGCTGACCTTTGGCTTGGCCAAAGCGAGG 1390
   |||||
Db 1415 CCTGCTCAGAGAGACCTGTGGCCAGGTCAAGCTGACCTTTGGCTTGGCCAAAGCGAGG 1474

QY 1391 GAAGGGGTAGACTCAAGCCGCTGCTCAGTAACTGAGAGCGGCCCGAGGCTCTCAACA 1450
   |||||
Db 1475 GAAGGGGTAGACTCAAGCCGCTGCTCAGTAACTGAGAGCGGCCCGAGGCTCTCAACA 1534

QY 1451 CGGGAAGTTCACAGCAAGTCCGAGTCTGAGATTTTGGGCTGCTCTGGAGAGCTCT 1510
   |||||
Db 1535 CGGGAAGTTCACAGCAAGTCCGAGTCTGAGATTTTGGGCTGCTCTGGAGAGCTCT 1594

QY 1511 CTCATATGACGAGGCTCCCTTACCCTTAAATGTCTACTGAAAGAGGTGCGAGGCTGGA 1570
   |||||
Db 1595 CTCATATGACGAGGCTCCCTTACCCTTAAATGTCTACTGAAAGAGGTGCGAGGCTGGA 1654

QY 1571 GAAGGGGTACCGGATGAGAACCCCGGAGGCTGTGCAAGGCCCGCTGACGCTCATGAG 1630
   |||||
Db 1655 GAAGGGGTACCGGATGAGAACCCCGGAGGCTGTGCAAGGCCCGCTGACGCTCATGAG 1714

QY 1631 CAGCTGCTGGAGGAGCAGACCCCGCCCGCCAGCCCTTCCGAAACTGGCGAGAGAGT 1690
   |||||
Db 1715 CAGCTGCTGGAGGAGCAGACCCCGCCCGCCAGCCCTTCCGAAACTGGCGAGAGAGT 1774

QY 1691 GGGCGGAGACTACGAGAGTGCAGAGTCCCGAGCTCTGCTCAAGGCGAGAGCGCGAGG 1750
   |||||
Db 1775 GGGCGGAGACTACGAGAGTGCAGAGTCCCGAGCTCTGCTCAAGGCGAGAGCGCGAGG 1834

QY 1751 CTCACACTGCGCCCGGAAGCCAGAGAGCTGACGCCCAACCCGAGGAGGCGCTTGGAGCC 1810
   |||||
Db 1835 CTCACACTGCGCCCGGAAGCCAGAGAGCTGACGCCCAACCCGAGGAGGCGCTTGGAGCC 1894

QY 1811 GGAACGAGAGTGAAGTGGCGCTGGGGGCACTGACAGGCCCAAGAGAGGCTCAGG 1870
   |||||
Db 1895 GGAACGAGAGTGAAGTGGCGCTGGGGGCACTGACAGGCCCAAGAGAGGCTCAGG 1954

QY 1871 CGGGCAAGTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1930
   |||||
Db 1955 CGGGCAAGTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2014

QY 1931 GCCCGGTGAGACACCCGAGACCTGCAAGATGATGCCCGGATTAAGAGCATTTCTAAG 1989
   |||||
Db 2015 GCCCGGTGAGACACCCGAGACCTGCAAGATGATGCCCGGATTAAGAGCATTTCTAAG 2073

RESULT 14
S71669 1744 bp mRNA linear PRI 24-OCT-1994
LOCUS S71669
DEFINITION Isk-leukocyte carboxyl-terminal src kinase related gene [human],
peripheral T cells, mRNA, 1744 nt.
ACCESSION S71669
VERSION S71669
KEYWORDS S71669.1 GI:559593
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1744)
AUTHORS McVicar,D.W., Lai,B.K., Lloyd,A., Kawamura,M., Chen,Y.Q., Zhang,X.,
Staples,J.E., Ortolano,J.R. and O'Shea,J.J.
Molecular cloning of Isk, a carboxyl-terminal src kinase (csk)
related gene, expressed in leukocytes
Oncogene 9 (7), 2037-2044 (1994)
JOURNAL MEDLINE
94268844

```

PUBMED 7516063
GenBank staff at the National Library of Medicine created this entry [NCBI gi15068] from the original journal article.

This sequence comes from Fig. 1.

FEATURES
Location/Qualifiers
1..1744

gene
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1744
/gene="lsk"
/note="leukocyte carboxyl-terminal src kinase related gene"

CDS
/note="leukocyte carboxyl-terminal src kinase related gene: This sequence comes from Fig. 1"
/protein_id="AB30995.1"
/db_xref="GI:559594"

BASE COUNT 354 a 534 c 582 g 274 t
ORIGIN

Query Match 82.6%; Score 1651.4; DB 9; Length 1744;
Best Local Similarity 98.3%; Pred. No. 3.8e-237;
Matches 1712; Conservative 0; Mismatches 21; Indels 8; Gaps 4;

260 GGGGGGGGAGGCTCTGTTCTGTTGGGGGAGGATTTACGAGCTGATTTCTGTAAGA 319
|||||
Db 4 GGGGGGGGAGGCTCTGTTCTGTTGGGGGAGGATTTACGAGCTGATTTCTGTAAGA 63
320 ACTTCCCGGGTGGAGCCCGCTTCTCCGAGCCTGGACACCCCGTCCGCTCAGCAG 379
|||||
Db 64 ACTTCCCGGGTGGAGCCCGCTTCTCCGAGCCTGGACACCCCGTCCGCTCAGCAG 123
380 GATGCCAAGAGGCGTGGGGCCCGGGGCAACAGTATTCACCAATGGAGCACACCG 439
|||||
Db 124 GATGCCAAGAGGCGTGGGGCCCGGGGCAACAGTATTCACCAATGGAGCACACCG 183
440 CCCCAGCCAGGAGGAGCTGGCTTCCGCAAGGAGGAGCTGATCCTCTGAGGCTG 499
|||||
Db 184 CCCCAGCCAGGAGGAGCTGGCTTCCGCAAGGAGGAGCTGATCCTCTGAGGCTG 243
500 CGAGAACAGAGCTGTACCGCGTCAAGACACACAGTGGAGGAGGAGGAGGAGTGG 559
|||||
Db 244 CGAGAACAGAGCTGTACCGCGTCAAGACACACAGTGGAGGAGGAGGAGGAGTGG 303
560 AGCTGGGGCGTGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 619
304 AGCTGGGGCGTGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 363
620 GTGGTTCACAGGAGAAATCTCGGGCCAGAGAGGCTGTCCAGACAGTGCAGCTCC 679
|||||
Db 364 GTGGTTCACAGGAGAAATCTCGGGCCAGAGAGGCTGTCCAGACAGTGCAGCTCC 423
680 TGGGCTGTCTGTGTCGGGAGAGTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 739
|||||
Db 424 TGGGCTGTCTGTGTCGGGAGAGTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 483
740 CTTTGGCCGAGAGCTATCAGTACCGCGTGTGTCAGCCGAGAGGAGGAGGAGGAGT 799
|||||
Db 484 GTTTGGCCGAGAGCTATCAGTACCGCGTGTGTCAGCCGAGAGGAGGAGGAGGAGT 543
800 TGAAGCCGCTGTTCTTCTGCAACCTATGACATGTGAGCATTTACAGCAGAGCAAG 859

|||||
Db 544 TGAAGCCGCTGTTCTTCTGCAACCTATGACATGTGAGCATTTACAGCAGAGCAAGG 603
860 CGTATCTGACCAACCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 919
604 CGCTATCTGACCAACCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 663
920 GAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 979
664 GAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 723
980 CGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1039
724 CGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
1040 GAAGATATCAAGTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1099
781 GAAGATATCAAGTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
1100 GAAGATATCAAGTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1159
841 GAAGATATCAAGTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
1160 CATTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1219
901 CATTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
1220 AGCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1279
961 AGCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
1280 GTACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1339
1021 GTACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
1340 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1399
1081 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
1400 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1459
1141 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
1460 CACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1519
1201 CACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
1520 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1579
1261 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
1580 CCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1639
1321 CCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
1640 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1699
1381 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
1700 GCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1759
1441 GCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
1760 GCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1819
1501 GCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1558
1820 GAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1879
1559 GAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1615
1880 CATCTCTGTTGTTCTTCTGCAACCTATGACATGTGAGCATTTACAGCAGAGCAAG 1939

Db 1616 CATCTCTCTGTGCTCCACAGCAGGAGGCTGCTCCACAGTATGAGGGGCTCTGGCGGCCCGCTGG 1675
QY 1940 ACACCTCCAGACCTCGGAGAGTATGATCGCCCATTAAGACGATTTCTAAGACATCTAATAA 1999
Db 1676 ACACCTCCAGACCTCGGAGAGTATGATCGCCCATTAAGAGGATTTCTAAGAGAAAAA 1735
QY 2000 A 2000
Db 1736 A 1786

RESULT 15
AR053291 1521 bp DNA Linear PAT 29-SEP-1999
LOCUS AR053291
DEFINITION Sequence 10 from patent US 5834208.
ACCESSION AR053291
VERSION AR053291.1 GI:5978153
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1521)
AUTHORS Sakano, S.
TITLE Tyrosine kinase
JOURNAL Patent: US 5834208-A 10 10-NOV-1998;
FEATURES
source 1. 1521
/organism="unknown"
BASE COUNT 295 a 470 c 507 g 249 t
ORIGIN

Query Match 76.0%; Score 1519.4; DB 6; Length 1521;
Best Local Similarity 99.9%; Pred. No. 2e-217;
Matches 1520; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 258 ATGGGGGGGCGAGGCTCTGTGTTCTGTGGGGGCAATTTCACGGCTGTGATTCTGCTAG 317
Db 1 ATGGGGGGGCGAGGCTCTGTGTTCTGTGGGGGCAATTTCACGGCTGTGATTCTGCTAG 60
QY 318 GAACCTCCCGGGGTAGAGCCCGGCTCTCGAGGCTGGACCCCGCTCCGCTCAGCC 377
Db 61 GAACCTCCCGGGGTAGAGCCCGGCTCTCGAGGCTGGACCCCGCTCCGCTCAGCC 120
QY 378 AGAATGCCAAGAGGCGCTGGGCGCCGGGCAACCGATATCAACCAATGGAGCACAC 437
Db 121 AGAATGCCAAGAGGCGCTGGGCGCCGGGCAACCGATATCAACCAATGGAGCACAC 180
QY 438 CGCCCAAGCAGAGGGAGCTGGCTTCCGCAAGGGGCACTGGTCAACATCTGAGGCC 497
Db 181 CGCCCAAGCAGAGGGAGCTGGCTTCCGCAAGGGGCACTGGTCAACATCTGAGGCC 240
QY 498 TGGCGAACAAGAGTGTATCCGCTCAAGACACCAACAGTGGAGAGAGGGCTGTG 557
Db 241 TGGCGAACAAGAGTGTATCCGCTCAAGACACCAACAGTGGAGAGAGGGCTGTG 300
QY 558 GCACCTGGGGGCTCGGGAGCGGAGGCGCTCTCCGAGACCCCAAGCTCAGCTCATG 617
Db 301 GCACCTGGGGGCTCGGGAGCGGAGGCGCTCTCCGAGACCCCAAGCTCAGCTCATG 360
QY 618 CCGTGTTCACAGGAGATCTGGGCGAAGAGGCTGTCCAGACCTCAGCTCCGAG 677
Db 361 CCGTGTTCACAGGAGATCTGGGCGAAGAGGCTGTCCAGACCTCAGCTCCGAG 420
QY 678 GATGGGCTGTCCTGTGGGAGTCCGCGGCGGCGGCGGCGGAGTACCTGTGGCGTG 737
Db 421 GATGGGCTGTCCTGTGGGAGTCCGCGGCGGCGGCGGCGGAGTACCTGTGGCGTG 480
QY 738 AGCTTTGGCGGAGAGTATCCATACGCGTGTGACACCGGAGGCGGCGGAGTATC 797
Db 481 AGCTTTGGCGGAGAGTATCCATACGCGTGTGACACCGGAGGCGGCGGAGTATC 540
QY 798 GATGAGGCGGCTTTCTTCTGCAACTCATGACATGTGGAGCATTTACAGCAAGCAAG 857

Db 541 GATGAGGCGGCTTTCTTCTGCAACTCATGACATGTGGAGCATTTACAGCAAGCAAG 600
QY 858 GGCCTATCTCTCACCACACTGTGTAGAACCAAGGGAACACGGGACCAATCTCGCCAG 917
Db 601 GGCCTATCTCTCACCACACTGTGTAGAACCAAGGGAACACGGGACCAATCTCGCCAG 660
QY 918 GAGGAGTGGCGGAGGCGGCTGTACTGAACTGACAGATTGTACATTGGAGACAG 977
Db 661 GAGGAGTGGCGGAGGCGGCTGTACTGAACTGACAGATTGTACATTGGAGACAG 720
QY 978 ATCGAGAGGAGATTGTGAGCTGTCTGCAAGGTGAGTACCTGGGGCAAAAGGTGCC 1037
Db 721 ATCGAGAGGAGATTGTGAGCTGTCTGCAAGGTGAGTACCTGGGGCAAAAGGTGCC 780
QY 1038 GTGAAGATATCACTGTGATGTACACCCAGGCTTCCGAGAGAGAGGCGGTCTAG 1097
Db 781 GTGAAGATATCACTGTGATGTACACCCAGGCTTCCGAGAGAGAGGCGGTCTAG 840
QY 1098 ACGAAGATGCAACACGAGAACCTGTGGCTCTCTGGGCGTGTATCTGACACAGGGCTG 1157
Db 841 ACGAAGATGCAACACGAGAACCTGTGGCTCTCTGGGCGTGTATCTGACACAGGGCTG 900
QY 1158 TACATTGTATGAGACACGTGAGCAAGGGCAACCTGTGTAACCTTTCGGAACCCGGGT 1217
Db 901 TACATTGTATGAGACACGTGAGCAAGGGCAACCTGTGTAACCTTTCGGAACCCGGGT 960
QY 1218 CGAGCCCTGTGAACACCGCTCAGCTCTGTGAGTTTCTGTGACGTGGCGGAGGGCTAG 1277
Db 961 CGAGCCCTGTGAACACCGCTCAGCTCTGTGAGTTTCTGTGACGTGGCGGAGGGCTAG 1020
QY 1278 GAGTACTGTGAGAGCAAGACCTGTGTGACCGGCACTGTGGCGGCGGCAACATCTGTGTC 1337
Db 1021 GAGTACTGTGAGAGCAAGACCTGTGTGACCGGCACTGTGGCGGCGGCAACATCTGTGTC 1080
QY 1338 TCAGAGGACCTGTGTGGCCAGAGTGTGGCTGTGGCCAAAGCCGAGCGGAGGGG 1397
Db 1081 TCAGAGGACCTGTGTGGCCAGAGTGTGGCTGTGGCCAAAGCCGAGCGGAGGGG 1140
QY 1398 CTAGACTCAAGCCGGCTGCGGTCAAGTGGAGGCGGCGGAGGCTCTCAACACGGGAG 1457
Db 1141 CTAGACTCAAGCCGGCTGCGGTCAAGTGGAGGCGGCGGAGGCTCTCAACACGGGAG 1200
QY 1458 TTCACACAGAGTGGATGTGTGAGTTTGGGGTGTGCTCTGTGGAGGCTTTCATAT 1517
Db 1201 TTCACACAGAGTGGATGTGTGAGTTTGGGGTGTGCTCTGTGGAGGCTTTCATAT 1260
QY 1518 GGAGCGGCTCCGTACCTTAATAATGTCACTGAAGAAGTGTGGAGGCCGTGAGAGGGG 1577
Db 1261 GGAGCGGCTCCGTACCTTAATAATGTCACTGAAGAAGTGTGGAGGCCGTGAGAGGGG 1320
QY 1578 TACCGCATGGAACCCCGGAGGCGTGTCCAGGCGGCGTGCAGCTCTCATGAGCAGTGC 1637
Db 1321 TACCGCATGGAACCCCGGAGGCGTGTCCAGGCGGCGTGCAGCTCTCATGAGCAGTGC 1380
QY 1638 TGGGAGGAGAGCCCGCGCGCGGCGCACCTTCCCAAACTGGCCGAGAGAGTGGCCGG 1697
Db 1381 TGGGAGGAGAGCCCGCGCGCGGCGCACCTTCCCAAACTGGCCGAGAGAGTGGCCGG 1440
QY 1698 GAGCTACGACATGCAAGTGTGCCACAGCTTCCTGTAGGGCAGAGAGCGGAGGCTCAC 1757
Db 1441 GAGCTACGACATGCAAGTGTGCCACAGCTTCCTGTAGGGCAGAGAGCGGAGGCTCAC 1500
QY 1758 TCGCCCGGAGGAGGAGGCC 1778
Db 1501 TCGCCCGGAGGAGGAGGCC 1521

Search completed: August 1, 2003, 18:16:28
Job time : 9498 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 15:19:49 ; Search time 182 seconds
(Without alignments) 4850.364 Million cell updates/sec

Title: US-09-977-260-1
Perfect score: 2000
Sequence: 1 ctgcctccaagtgtgtcagc.....attctaagactcttaaaaa 2000

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapept 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfillseq.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2000	100.0	2000	US-08-426-509A-1	Sequence 1, Appli
2	2000	100.0	2000	US-08-232-545-1	Sequence 1, Appli
3	2000	100.0	2000	PCT-US95-05008-1	Sequence 1, Appli
4	1909.8	95.5	1987	US-08-876-882-1	Sequence 1, Appli
5	1909.8	95.5	1987	US-09-315-928-1	Sequence 1, Appli
6	1903.4	95.2	1942	US-08-604-989A-11	Sequence 1, Appli
7	1519.4	76.0	1521	US-08-604-989A-10	Sequence 10, Appli
8	1398	69.9	1398	US-08-604-989A-9	Sequence 9, Appli
9	1377	73.8	1713	US-09-741-154-1	Sequence 1, Appli
10	738	36.9	738	US-08-604-989A-8	Sequence 8, Appli
11	455	22.8	16389	US-09-741-154-3	Sequence 3, Appli
12	225	11.2	225	US-08-604-989A-7	Sequence 7, Appli
13	212.4	10.6	1611	US-07-820-011A-3	Sequence 3, Appli
14	212.4	10.6	1611	PCT-US93-00445-3	Sequence 3, Appli
15	199.2	10.0	1602	US-07-820-011A-1	Sequence 1, Appli
16	199.2	10.0	1602	PCT-US93-00445-1	Sequence 1, Appli
17	192	9.6	192	US-08-604-989A-6	Sequence 6, Appli
18	185	9.2	1574	US-09-173-581-12	Sequence 12, Appli
19	185	9.2	1574	US-09-420-915-12	Sequence 12, Appli
20	181.8	9.1	3623	US-08-306-691B-35	Sequence 35, Appli
21	170.4	8.5	1467	US-09-579-182-2	Sequence 2, Appli
22	168.8	8.4	1548	US-09-099-053-1	Sequence 1, Appli
23	165.8	8.3	2674	US-09-817-180-1	Sequence 1, Appli
24	162	8.1	5993	US-09-383-630-1	Sequence 1, Appli
25	162	8.1	5993	US-09-383-630-2	Sequence 2, Appli
26	151.2	7.6	2647	US-09-220-132-77	Sequence 77, Appli
27	151.2	7.6	2647	PCT-US93-06251-77	Sequence 77, Appli

28	148.8	7.4	2049	US-09-099-749-10	Sequence 10, Appli
29	148.8	7.4	2433	US-09-620-312D-830	Sequence 830, App
30	148.8	7.4	2598	US-09-417-197-110	Sequence 110, App
31	148.8	7.4	2616	US-09-417-197-108	Sequence 108, App
32	148	7.4	2469	US-08-459-286-1	Sequence 1, Appli
33	146.4	7.3	933	US-08-701-191A-4	Sequence 4, Appli
34	146.4	7.3	1056	US-08-701-191A-5	Sequence 5, Appli
35	146.4	7.3	2469	US-07-997-133-2	Sequence 2, Appli
36	146.4	7.3	2469	US-07-997-133-2	Sequence 2, Appli
37	146.4	7.3	2662	US-08-451-822A-14	Sequence 14, Appli
38	146.4	7.3	2662	US-08-323-430-14	Sequence 14, Appli
39	146.4	7.3	2733	US-08-371-001-14	Sequence 14, Appli
40	146.4	7.3	2733	PCT-US96-00331-14	Sequence 14, Appli
41	137.6	6.9	1491	US-09-006-675-1	Sequence 1, Appli
42	137.6	6.9	1491	US-09-228-603A-1	Sequence 1, Appli
43	136.6	6.8	3945	US-09-016-434-1404	Sequence 1404, Ap
44	136.6	6.8	3969	US-08-436-044-5	Sequence 5, Appli
45	136.6	6.8	3969	US-08-222-616-23	Sequence 23, Appli

ALIGNMENTS

RESULT 1
US-08-426-509A-1
Sequence 1, Application US/08426509A

Patent No. 6326469

GENERAL INFORMATION:

APPLICANT: Gillich, Axel

APPLICANT: Gishitzky, Mikhail

APPLICANT: Sures, Irman G.

TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN

TITLE OF INVENTION: TYROSINE KINASES

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York,

STATE: NY

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/426,509A

FILING DATE: 21-APR-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/232,545

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A

REGISTRATION NUMBER: 7683-0074-999

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-9741

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2000 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

US-08-426-509A-1

Query Match 100.0%; Score 2000; DB 4; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	1	CTCGCTCCAAAGTTGTCAGACCGGGACCGCTCTCGGGGTGTGCAGCCGGCTCGGGGAGGCC	60
Db	1	CTCGCTCCAAAGTTGTGTGCAGCCGGACCGCTCTCGGGGTGTGCAGCCGGCTCGGGGAGGCC	60
OY	61	TCCTGGGGGCGGGGCGCGGGGGCGGCTCGGGGGGCGCCCTCTGAGCAGAAAAACAGGAAGAAC	120
Db	61	TCCTGGGGGCGGGGCGCGGGGGCGGCTCGGGGGGCGCCCTCTGAGCAGAAAAACAGGAAGAAC	120
OY	121	AGGCTCGGTTCCAGTGTGGACCCAGCTTCCTACCTCTGTGTCCAGCGCGCTGGCTGTGGCA	180
Db	121	AGGCTCGGTTCCAGTGTGGACCCAGCGCTTCCTCTGTGTCCAGCGCGCTGGCTGTGGCA	180
OY	181	GGCCATTTCCACAGCGTCCCGGAGCTGTGACACTTGTCAAGTGGCTCAGCTCACCTGGCTCAG	240
Db	181	GGCCATTTCCACAGCGTCCCGGAGCTGTGACCACTTGTCTAAGTGTGCTCTCAGCTGGCTCAG	240
OY	241	TTTCCCTCTGGGGGGGCGATGGCGGGGCGAGGCTCTCTGTGTTCTCTGCGGGGCAATTTACG	300
Db	241	TTTCCCTCTGGGGGGGCGATGGCGGGGCGAGGCTCTGTGTGTTCTCTGCGGGGCAATTTACG	300
OY	301	GCTGTGATTTCTCTAGAGAACTTCCCGGGGTGAGCCCGGCTTCTCGAGCTGTGGAC	360
Db	301	GCTGTGATTTCTCTAGAGAACTTCCCGGGGTGAGCCCGGCTTCTCTCGAGCTGTGGAC	360
OY	361	CCCCCCCCGTCACGCCAGAGATGCCACAGAGGCGCTGGGCCCCCGGGGACCCAGTGTATCA	420
Db	361	CCCCCCCCGTCACGCCAGAGATGCCACAGAGGCGCTGGGCCCCCGGGGACCCAGTGTATCA	420
OY	421	CCAAATGCGAGCACACCCCGCCCAAGGCCAGGGGAGCTGGCCTTCGCGAAGGGCGAGTGG	480
Db	421	CCAAATGCGAGCACACCCCGCCCAAGGCCAGGGGAGCTGGCCTTCGCGAAGGGCGAGTGG	480
OY	481	TCACCATCTGTGAGGCTGTGCAGAAACAAGCTGGTACCGCGTCAAGACCAACACACAGTGG	540
Db	481	TCACCATCTGTGAGGCTGTGCAGAAACAAGCTGGTACCGCGTCAAGACCAACACACAGTGG	540
OY	541	GACAGGAGGGGCTGTGTGGACCTGGGGGGGCTGCGGGAGCGGGAGCGCCCTTCCGAGAC	600
Db	541	GACAGGAGGGGCTGTGTGGACCTGGGGGGGCTGCGGGAGCGGGAGCGCCCTTCCGAGAC	600
OY	601	CCAAAGCTCAGGCTCATTCGCGTGTGTCCACGGGGAAMATTCGGGGCAGGAGGCTGTCCAGC	660
Db	601	CCAAAGCTCAGGCTCATTCGCGTGTGTCCACGGGGAAMATTCGGGGCAGGAGGCTGTCCAGC	660
OY	661	AGCTGTACGCTTCCGAGGATGGGCTGTTCTGTGTGTCGGGAGTCCGCGGCCACCCCGGCG	720
Db	661	AGCTGTACGCTTCCGAGGATGGGCTGTTCTGTGTGTCGGGAGTCCGCGGCCACCCCGGCG	720
OY	721	ACTACGTCTGTGGGTGTGAGCTTTGGCGCGGAGCTCATTCACCTACCGGCTGTGCACCGCG	780
Db	721	ACTACGTCTGTGGGTGTGAGCTTTGGCGCGGAGCTCATTCACCTACCGGCTGTGCACCGCG	780
OY	781	ACGGCACCTTACATCATGTAGAGCGCGTGTCTTCTGTCAACCTCATGTGACATGTGTGAGC	840
Db	781	ACGGCACCTTACATCATGTAGAGCGCGTGTCTTCTGTCAACCTCATGTGACATGTGTGAGC	840
OY	841	ATTACAGCAGACAAAGGGCGCTATCTGTACACCAAGCTGTGTGAGACCAAAAGGGGAAACAG	900
Db	841	ATTACAGCAGACAAAGGGCGCTATCTGTACACCAAGCTGTGTGAGACCAAAAGGGGAAACAG	900
OY	901	GGACCAATCGGGCGGAGAGAGAGCTGGCCAGAGGGGGCTGGTACTGACACTGCAGCATTT	960
Db	901	GGACCAATCGGGCGGAGAGAGAGCTGGCCAGAGGGGGCTGGTACTGACACTGCAGCATTT	960
OY	961	TGACATTGTGGAGCACAGATCGGAGGGGAGGTTTGGAGCGTCTCGTAGGGGTAGATACC	1020
Db	961	TGACATTGTGGAGCACAGATCGGAGGGGAGGTTTGGAGCGTCTCGTAGGGGTAGATACC	1020
OY	1021	TGGGGCAAAAGTGTGGCGTGAAGAAATATCAAGTGTGATGTACACGCCAGGCTTCTCTGG	1080
Db	1021	TGGGGCAAAAGTGTGGCGTGAAGAAATATCAAGTGTGATGTACACGCCAGGCTTCTCTGG	1080
OY	1081	ACGAGAGCGCGCTATGTACGAAGATTCACACGAGAACCTGTGTGCGTCTCTGGGCGTGA	1140

Db	1081	ACGAGACGGCGTCAATGACGAGATGCAACACGAGAACCTGTGTGCTCTCGGGCGTGA	1140
QY	1141	TCCTGACACGAGGGGCTGTACATTTGTCATGAGACACGTAGGCAAGGGCAACCTGGTGAAT	1200
Db	1141	TCCTGACACGAGGGGCTGTACATTTGTCATGAGACACGTAGGCAAGGGCAACCTGGTGAAT	1200
QY	1201	TTCTGCGGACCCGGGGCTGAGCCCTCGTGAACACCGCTCAGCTCTGCAATTTCTCTGC	1260
Db	1201	TTCTGCGGACCCGGGGCTGAGCCCTCGTGAACACCGCTCAGCTCTGCAATTTCTCTGC	1260
QY	1261	ACGTGGCCGAGGGCATGTGAGTACCTGGAGACCAAGAACCTTGTGACCGGACCTGGCGG	1320
Db	1261	ACGTGGCCGAGGGCATGTGAGTACCTGGAGACCAAGAACCTTGTGACCGGACCTGGCGG	1320
QY	1321	CCCGCAACATCCTGGTCTCAGAGGACCTGTGTGGCCAAAGTCAAGGCACTTTGGCTGGCCA	1380
Db	1321	CCCGCAACATCCTGGTCTCAGAGGACCTGTGTGGCCAAAGTCAAGGCACTTTGGCTGGCCA	1380
QY	1381	AAGCCGAGCGGAAGGGGCTAGACTCAAGCCGGCTGCCCCGTAAGTGAAGCGGCCCGCAGG	1440
Db	1381	AAGCCGAGCGGAAGGGGCTAGACTCAAGCCGGCTGCCCCGTAAGTGAAGCGGCCCGCAGG	1440
QY	1441	CTCTCAACACGGGGAGTTTCAACACGACAGTGGAGTGTGAGATTTTGGGGTGGCGTCT	1500
Db	1441	CTCTCAACACGGGGAGTTTCAACACGACAGTGGAGTGTGAGATTTTGGGGTGGCGTCT	1500
QY	1501	GGGAGGCTTCTCATATATGAGACGGGGCTCCGTACCCCTAAATATCTCACTGAAGAGTGTGCG	1560
Db	1501	GGGAGGCTTCTCATATATGAGACGGGGCTCCGTACCCCTAAATATCTCACTGAAGAGTGTGCG	1560
QY	1561	AGGCCGTGGAGAGGGGCTACCGCATGGAACCCCCCGAGGGCTGTCCAGGGCCCCGTGCACG	1620
Db	1561	AGGCCGTGGAGAGGGGCTACCGCATGGAACCCCCCGAGGGCTGTCCAGGGCCCCGTGCACG	1620
QY	1621	TCCTCATGAGAGCTGCTGGGAGGACAGACCCCGCCGGCCCAACCTTCGGCAAACTGG	1680
Db	1621	TCCTCATGAGAGCTGCTGGGAGGACAGACCCCGCCGGCCCAACCTTCGGCAAACTGG	1680
QY	1681	CCGAGAACTGGACCCGGGAGCTACGCAATGCAAGTGCACAGCCTCCGTCACAGGCGAGG	1740
Db	1681	CCGAGAACTGGACCCGGGAGCTACGCAATGCAAGTGCACAGCCTCCGTCACAGGCGAGG	1740
QY	1741	ACGCCAGCGGCTCCACCTCGCCCGGAAAGCCAGAGACCTTGACCCACCCGGTGGGGGCTT	1800
Db	1741	ACGCCAGCGGCTCCACCTCGCCCGGAAAGCCAGAGACCTTGACCCACCCGGTGGGGGCTT	1800
QY	1801	TGGCCCAAGAGACCGAGAGATGAGATGAGATGCGGCGTGGGGGCACTCAACAGGCGCCAAAG	1860
Db	1801	TGGCCCAAGAGACCGAGAGATGAGATGAGATGCGGCGTGGGGGCACTCAACAGGCGCCAAAG	1860
QY	1861	AGGGTCCAGCGGGGCAAGTCAATCTCGCTGGTGCACAGCAGGAGGGCTGGCCACAGTAGAGG	1920
Db	1861	AGGGTCCAGCGGGGCAAGTCAATCTCGCTGGTGCACAGCAGGAGGGCTGGCCACAGTAGAGG	1920
QY	1921	GGCTGTGGGGCGGCCGTGGAGACACCCAGACCTCGGAAGAGATGATGCCCGGTAAGAAGCG	1980
Db	1921	GGCTGTGGGGCGGCCGTGGAGACACCCAGACCTCGGAAGAGATGATGCCCGGTAAGAAGCG	1980
QY	1981	ATTCTAAGAGACTTAAAAA 2000	
Db	1981	ATTCTAAGAGACTTAAAAA 2000	

TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,545
FILING DATE: 22-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2000 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-232-545-1

Query Match 100.0%; Score 2000; DB 4; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGCTCCAGTTGTGACGCGGCGGCTCTCGGGGTGTGACGCGCGCTCGCGAGGCGCC 60
DB 1 CTCGCTCCAGTTGTGACGCGGCGGCTCTCGGGGTGTGACGCGCGCTCGCGAGGCGCC 60
QY 61 TCCTGGGGGCGGGCGGGGCGGCTCGGGGGCGCCCTGAGCAGAAAAAGAGAAACC 120
DB 61 TCCTGGGGGCGGGCGGGGCGGCTCGGGGGCGCCCTGAGCAGAAAAAGAGAAACC 120
QY 121 AGGCTCGGTCCAGTGGACCCAGCTCCTTACCTCTGTGACAGCGCGCTGGCTGTGCA 180
DB 121 AGGCTCGGTCCAGTGGACCCAGCTCCTTACCTCTGTGACAGCGCGCTGGCTGTGCA 180
QY 181 GGCATTCGCAAGCGCCCGCACTGTGACACTGTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 181 GGCATTCGCAAGCGCCCGCACTGTGACACTGTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 241 TTTCCCTCTGGGGGCGATGGCGGGGCGAGGCTCTGTGTTCTGCGGGCATTTACG 300
DB 241 TTTCCCTCTGGGGGCGATGGCGGGGCGAGGCTCTGTGTTCTGCGGGCATTTACG 300
QY 301 GCTGTGATTTGCTGAGGAACCTTCCCGGGGTAGCGCCCGCTTCTCCGAGCTGGCACC 360
DB 301 GCTGTGATTTGCTGAGGAACCTTCCCGGGGTAGCGCCCGCTTCTCCGAGCTGGCACC 360
QY 361 CCCCGCCGTCGACGCGAGATGCCAAGAGCGCTGGGCGCCGGGCGACCAAGTATCA 420
DB 361 CCCCGCCGTCGACGCGAGATGCCAAGAGCGCTGGGCGCCGGGCGACCAAGTATCA 420
QY 421 CCAATGCGACACACCGCCCAAGCCAGAGGAGACTGGCTTCCGCAAGGGCGAGCTGG 480
DB 421 CCAATGCGACACACCGCCCAAGCCAGAGGAGACTGGCTTCCGCAAGGGCGAGCTGG 480
QY 481 TCACCATCTGGAGGCTGCGAAGCAAGAGCTGTACCGCGTCAAGCACCACACAGTG 540
DB 481 TCACCATCTGGAGGCTGCGAAGCAAGAGCTGTACCGCGTCAAGCACCACACAGTG 540

DB 481 TCACCATCTGGAGGCTGCGAAGCAAGAGCTGTACCGCGTCAAGCACCACACAGTG 540
QY 541 GACAGAGGGGCTGTGACACTGGGGGCTGCGGGAGCGGAGGCGCTTCTCCGAGACC 600
DB 541 GACAGAGGGGCTGTGACACTGGGGGCTGCGGGAGCGGAGGCGCTTCTCCGAGACC 600
QY 601 CCAAGCTAGGCTCATGCGCGGTGTCCACGGGAAGATCTCGGGCCAGAGAGCTGTCCAGC 660
DB 601 CCAAGCTAGGCTCATGCGCGGTGTCCACGGGAAGATCTCGGGCCAGAGAGCTGTCCAGC 660
QY 661 AGCTGACGCTCCGAGGATGGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 661 AGCTGACGCTCCGAGGATGGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 ACTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB 721 ACTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 781 ACGGCACTCACAATCATGATGAGCGCGTCTTCTGCAACCTCATGAGCATGTGTGAGC 840
DB 781 ACGGCACTCACAATCATGATGAGCGCGTCTTCTGCAACCTCATGAGCATGTGTGAGC 840
QY 841 ATTACGACAGACAAAGGGCGCTATCTGACCAACGCTGTGAGACCAAGCGGAAACAG 900
DB 841 ATTACGACAGACAAAGGGCGCTATCTGACCAACGCTGTGAGACCAAGCGGAAACAG 900
QY 901 GGACCAATCGCGCGAGAGAGAGCTGGCGGAGGGGCTGTGTTACTGAACTGTGACGATT 960
DB 901 GGACCAATCGCGCGAGAGAGAGCTGGCGGAGGGGCTGTGTTACTGAACTGTGACGATT 960
QY 961 TGACATTTGGAGACAGATCGGAGAGAGAGATTGGAGCTGTCTGCAAGGCTGAGTACC 1020
DB 961 TGACATTTGGAGACAGATCGGAGAGAGAGATTGGAGCTGTCTGCAAGGCTGAGTACC 1020
QY 1021 TGGGCAAAAGGTGGCGCTGAAGAAATATCAAGTGTGATGTGACACCGACCTTCTGCG 1080
DB 1021 TGGGCAAAAGGTGGCGCTGAAGAAATATCAAGTGTGATGTGACACCGACCTTCTGCG 1080
QY 1081 ACGAGACGGCGCTCATGACGAAGATGCAACACAGAACTGTGCTGCTGCTGCTGCTG 1140
DB 1081 ACGAGACGGCGCTCATGACGAAGATGCAACACAGAACTGTGCTGCTGCTGCTGCTG 1140
QY 1141 TCCTGACACAGGGGCTGTACATTTGTATGAGACAGTGTGACAGGCGAACTGTGTA 1200
DB 1141 TCCTGACACAGGGGCTGTACATTTGTATGAGACAGTGTGACAGGCGAACTGTGTA 1200
QY 1201 TTTCTGGGACCGGGGGCTGAGCCCTGTGAACACCGCTGACCTCTGCAATTTCTGTC 1260
DB 1201 TTTCTGGGACCGGGGGCTGAGCCCTGTGAACACCGCTGACCTCTGCAATTTCTGTC 1260
QY 1261 ACGTGGCGGAGGGCATGAGTATCGTGGAGACCAAGAACTGTGACCGGCACTGGGCG 1320
DB 1261 ACGTGGCGGAGGGCATGAGTATCGTGGAGACCAAGAACTGTGACCGGCACTGGGCG 1320
QY 1321 CCCGCAACATCTGCTCTCAGAGACCTGTGTGGCCAAAGTATGAGCATTTGGCTGGGCA 1380
DB 1321 CCCGCAACATCTGCTCTCAGAGACCTGTGTGGCCAAAGTATGAGCATTTGGCTGGGCA 1380
QY 1381 AAGCCGAGCGGAAGGGCTAGACTCAACCGGCTGCGCTGAAGTGACAGCGCGCCGAG 1440
DB 1381 AAGCCGAGCGGAAGGGCTAGACTCAACCGGCTGCGCTGAAGTGACAGCGCGCCGAG 1440
QY 1441 CTCCTCAACACGGGAAGTTCACACAGAGTGGAGTGTGAGTTTGGGGTGGCTGCT 1500
DB 1441 CTCCTCAACACGGGAAGTTCACACAGAGTGGAGTGTGAGTTTGGGGTGGCTGCT 1500
QY 1501 GGGAGGCTTCTCATATGAGAGGGCTCCGTACCTTAATATTCATGAAAGAGTGTGG 1560
DB 1501 GGGAGGCTTCTCATATGAGAGGGCTCCGTACCTTAATATTCATGAAAGAGTGTGG 1560
QY 1561 AGGCGGTGAGAGAGGGGTACCGCATGGAACCCCGGAGGCTGTCCAGGCGCCGTCACG 1620
DB 1561 AGGCGGTGAGAGAGGGGTACCGCATGGAACCCCGGAGGCTGTCCAGGCGCCGTCACG 1620

QY 901 GGACCAAGTCGGCCGAGAGAGAGCTGGCCAGGGGGCTGGTTACTGAACCTGCAGCATT 960
DB 901 GGACCAAGTCGGCCGAGAGAGAGCTGGCCAGGGGGCTGGTTACTGAACCTGCAGCATT 960
QY 961 TGCATTTGGGAGCACAATCGGAGAGAGAGAGTTGGAGCTGTCTCTGAGGGTGAATACC 1020
DB 961 TGCATTTGGGAGCACAATCGGAGAGAGAGAGTTGGAGCTGTCTCTGAGGGTGAATACC 1020
QY 1021 TGGGGCAAAAGTGGCCGTGAGAGAAATCAAGTGTGATGTGACAGCCAGGCTTCTCTGG 1080
DB 1021 TGGGGCAAAAGTGGCCGTGAGAGAAATCAAGTGTGATGTGACAGCCAGGCTTCTCTGG 1080
QY 1081 ACAGAGAGCGCCGTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 1081 ACAGAGAGCGCCGTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 TCCATGACCAAGGGCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 TCCATGACCAAGGGCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 TTTGCGGAGCCCGGGGGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1201 TTTGCGGAGCCCGGGGGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1261 ACATGCGGAG 1320
DB 1261 ACATGCGGAG 1320
QY 1321 CCGGCAACATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
DB 1321 CCGGCAACATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1381 AAGCGAG 1440
DB 1381 AAGCGAG 1440
QY 1441 CTCTCAAAACAGGGAG 1500
DB 1441 CTCTCAAAACAGGGAG 1500
QY 1501 GGGAGGCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1501 GGGAGGCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
QY 1561 AGGCGCTGAG 1620
DB 1561 AGGCGCTGAG 1620
QY 1621 TCCATGAG 1680
DB 1621 TCCATGAG 1680
QY 1681 CCGAG 1740
DB 1681 CCGAG 1740
QY 1741 ACAGCGAG 1800
DB 1741 ACAGCGAG 1800
QY 1801 TGGCGCCAG 1860
DB 1801 TGGCGCCAG 1860
QY 1861 AGGCTTCAG 1920
DB 1861 AGGCTTCAG 1920
QY 1921 GGGCTTGGGAG 1980
DB 1921 GGGCTTGGGAG 1980
QY 1981 ATTCTAAGAGACTTAATAAAA 2000

DB 1981 ATTCTAAGAGACTTAATAAAA 2000

RESULT 4
US-08-876-882-1
Sequence 1, Application US/08876882
Patent No. 5981201
GENERAL INFORMATION:
APPLICANT: Avraham, Hava
APPLICANT: Groopman, Jerome E.
TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT
TITLE OF INVENTION: OF BREAST CANCER
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/876,882
FILING DATE: 16-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/035,228
FILING DATE: 08-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Doreen, Hogle M
REGISTRATION NUMBER: 36,361
REFERENCE/DOCKET NUMBER: NEDH97-01pa
TELECOMMUNICATION INFORMATION:
TELEPHONE: 781-861-6240
TELEFAX: 781-861-9540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1987 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-876-882-1

Query Match 95.5%; Score 1909.8; DB 2; Length 1987;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1973; Conservative 0; Mismatches 7; Indels 9; Gaps 5;

QY 1 CTCGCTTCCAGATTGTGAG 60
DB 8 CTCGCTTCCAGATTGTGAG 67
QY 61 TCTTGG 120
DB 68 TCTTGG 127
QY 121 AGGCTGCTCAG 180
DB 128 AGGCTGCTCAG 187
QY 181 GGGCATTTCCAG 240
DB 188 GGGCATTTCCAG 247
QY 241 TTTCCCTCTGG 300
DB 248 TTTCC--TCTGG 305

QY 301 GGTGATTTCTGTGAGAACTTCCCGGGGTAGACCCCGCTTCTCGAGCCTGGCACC 360
 DB 306 GGTGATTTCTGTGAGAACTTCCCGGGGTAGACCCCGCTTCTCGAGCCTGGCACC 365
 QY 361 CCCCTCCCGTCTGAGCAGATGCCAAGAGGGGCGGGGGCCCGGGCACCAGTGTATCA 420
 DB 366 CCCCTCCCGTCTGAGCAGATGCCAAGAGGGGCGGGGGCCCGGGCACCAGTGTATCA 425
 QY 421 CCAAAATGCGAGACACCCGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGGGCGAGTGG 480
 DB 426 CCAAAATGCGAGACACCCGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGGGCGAGTGG 485
 QY 481 TCACCATCTGTGAGAGCCTTCCGAGAACAGAGCTGGTACCGCTCAAGACCAACAGTGG 540
 DB 486 TCACCATCTGTGAGAGCCTTCCGAGAACAGAGCTGGTACCGCTCAAGACCAACAGTGG 545
 QY 541 GACAGAGAGGGGCTGCTGGAGAGCTGGGGGCGCTGGGGAGGCGGGGCGCTTCCGAGAC 600
 DB 546 GACAGAGAGGGGCTGCTGGAGAGCTGGGGGCGCTGGGGAGGCGGGGCGCTTCCGAGAC 605
 QY 601 CCAAGCTCAGCTCATGCGGTGGTTCACAGGGAGATCTCGGGCCAGAGAGCTGTCCAC 660
 DB 606 CCAAGCTCAGCTCATGCGGTGGTTCACAGGGAGATCTCGGGCCAGAGAGCTGTCCAC 665
 QY 661 AGCTGAGCCTCCGAGAGATGGGCTGTCTGCTGGGAGTCCGCGCCACCCGCGG 720
 DB 666 AGCTGAGCCTCCGAGAGATGGGCTGTCTGCTGGGAGTCCGCGCGCCACCCGCGG 725
 QY 721 ACTAGCTCTGTGCGTGTGAGCTTTGGCGCGGAGCTCATCCGCTGCTGACCGCG 780
 DB 726 ACTAGCTCTGTGCGTGTGAGCTTTGGCGCGGAGCTCATCCGCTGCTGACCGCG 785
 QY 781 AGGCGCACCTCACAATCGATGAGGCGGCTTCTTGCACCACTCATGAGCATGTGGAGC 840
 DB 786 AGGCGCACCTCACAATCGATGAGGCGGCTTCTTGCACCACTCATGAGCATGTGGAGC 845
 QY 841 ATTACAGCAAGACAGAGGGCGCTATCTGACCAAGCTGGTGGAGACCAAGGCGGAAACAG 900
 DB 846 ATTACAGCAAGACAGAGGGCGCTATCTGACCAAGCTGGTGGAGACCAAGGCGGAAACAG 905
 QY 901 GGACCAAGTGGCGCGGAGAGAGCTGGCCAGGGGCGGTACTGAACTGACGACTT 960
 DB 906 GGACCAAGTGGCGCGGAGAGAGCTGGCCAGGGGCGGTACTGAACTGACGACTT 965
 QY 961 TGACATTGGGAGCACAGATCGAGAGGAGAGTGGAGCTGCTGACGGGTGAGTACC 1020
 DB 966 TGACATTGGGAGCACAGATCGAGAGGAGAGTGGAGCTGCTGACGGGTGAGTACC 1025
 QY 1021 TGGGCGCAAAAGGTGGCGCTGAGAAATATCAATGATGTGACAGCCGAGGCTTCTGG 1080
 DB 1026 TGGGCGCAAAAGGTGGCGCTGAGAAATATCAATGATGTGACAGCCGAGGCTTCTGG 1085
 QY 1081 AGGAGACGGCGCTCATGAGAGATGCAACAGAGAACTGGTGGCTCTCTGGGCGTGA 1140
 DB 1086 AGGAGACGGCGCTCATGAGAGATGCAACAGAGAACTGGTGGCTCTCTGGGCGTGA 1145
 QY 1141 TCCTGACACAGGGGCTGTACATTTGATGAGACAGCTGAGCAAGGGCAACTGGTGA 1200
 DB 1146 TCCTGACACAGGGGCTGTACATTTGATGAGACAGCTGAGCAAGGGCAACTGGTGA 1205
 QY 1201 TTCTGCGGAGCGGGGGTGGAGCGCTGCAACACCGCTAGCTCTCGAGTTTCTCTGC 1260
 DB 1206 TTCTGCGGAGCGGGGGTGGAGCGCTGCAACACCGCTAGCTCTCGAGTTTCTCTGC 1265
 QY 1261 AGTGGCCGAGGAGCATGAGTACTGTGAGAGCAAGAGTGTGACACCGAGCTGGCCG 1320
 DB 1266 AGTGGCCGAGGAGCATGAGTACTGTGAGAGCAAGAGTGTGACACCGAGCTGGCCG 1325
 QY 1321 CCCGCAACATCTGTGTCAGAGGACCTGGTGGCCAGAGTCAAGCTTTGGCTGGCCA 1380
 DB 1326 CCCGCAACATCTGTGTCAGAGGACCTGGTGGCCAGAGTCAAGCTTTGGCTGGCCA 1385
 QY 1381 AAGCGAGCGGAAAGGGGCTAGACTCAAGCGGCTGCCCTCAAGTGGAGCGGCCCGCAGAG 1440

DB 1386 AAGCGAGCGGAAAGGGGCTAGACTCAAGCGGCTGCGCTCAAGTGGAGCGGCCGAGAG 1445
 QY 1441 CTCTCAACACGGAAGTTTCACAGCAAGTGGAGTGTGAGATTGAGGCTGTCTCT 1500
 DB 1446 CTCTCAACACGGAAGTTTCACAGCAAGTGGAGTGTGAGATTGAGGCTGTCTCT 1502
 QY 1501 GGGAGGCTTCTCATATGAGAGCGGCTCCGTACCTTAAATGTCTAGTAAAGAGTGTGCG 1560
 DB 1503 GGGAGGCTTCTCATATGAGAGCGGCTCCGTACCTTAAATGTCTAGTAAAGAGTGTGCG 1562
 QY 1561 AGGCGGTGAGAAAGGGGTACCGCATGTGAACCCCGAGGGCTGTGAGGCCCGCTGACAG 1620
 DB 1563 AGGCGGTGAGAAAGGGGTACCGCATGTGAACCCCGAGGGCTGTGAGGCCCGCTGACAG 1622
 QY 1621 TCCTCATGAGAGCTGTGAGAGGAGAGCCCGCGCGGACCGCTTCCGCAAACTGG 1680
 DB 1623 TCCTCATGAGAGCTGTGAGAGGAGAGCCCGCGCGGACCGCTTCCGCAAACTGG 1681
 QY 1681 CCGAGAGCTGTGAGCGGAGGCTAGAGGAGTGTGAGGCTTCCGCTCTGAGGCGAG 1740
 DB 1682 CCGAGAGCTGTGAGCGGAGGCTAGAGGAGTGTGAGGCTTCCGCTCTGAGGCGAG 1741
 QY 1741 AGCGGAGCGGCTTCACACTGCGCCGAGAGCGCTGACCGCCACCGGCTGGGGCGCT 1800
 DB 1742 AGCGGAGCGGCTTCACACTGCGCCGAGAGCGGCTGACCGCCACCGGCTGGGGCGCT 1798
 QY 1801 TGGCCCGAGAGAGCGAGAGTGTGAGAGTGTGAGGCTGGGCGGCTGACAGAGCGCCAAAG 1860
 DB 1799 TGGCCCGAGAGAGCGAGAGTGTGAGAGTGTGAGGCTGGGCGGCTGACAGAGCGCCAAAG 1858
 QY 1861 AGGCTCAGAGCGGAGAGTGTGAGAGTGTGAGGCTGGGCGGCTGACAGAGCGCCAAAG 1920
 DB 1859 AGGCTCAGAGCGGAGAGTGTGAGAGTGTGAGGCTGGGCGGCTGACAGAGCGCCAAAG 1918
 QY 1921 GGCTCTGGGCGGCGGCTGAGACCGCCAGAGCGGCTGAGAGTGTGAGGCTGGGCGGCT 1980
 DB 1919 GGCTCTGGGCGGCGGCTGAGACCGCCAGAGCGGCTGAGAGTGTGAGGCTGGGCGGCT 1978
 QY 1981 ATTCTAAGG 1989
 DB 1979 ATTCTAAGG 1987

RESULT 5
 US-09-315-928-1
 ; Sequence 1, Application US/09315928
 ; Patent No. 6368796
 ; GENERAL INFORMATION:
 ; APPLICANT: Avraham, Hava
 ; TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT OF
 ; FILE REFERENCE: NED97-01PAZ
 ; CURRENT APPLICATION NUMBER: US/09/315,928
 ; CURRENT FILING DATE: 1999-05-20
 ; PRIOR APPLICATION NUMBER: US 08/876,882
 ; PRIOR FILING DATE: 1997-06-16
 ; PRIOR APPLICATION NUMBER: US 60/035,228
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FASTSEQ for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 1987
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (263) ... (1846)
 US-09-315-928-1

Query Match 95.5%; Score 1909.8; DB 4; Length 1987;
 Best Local Similarity 99.2%; Pred. No. 0;

Matches 1973; Conservative 0; Mismatches 7; Indels 9; Gaps 5;

[illegible]

QY	1081	ACGAGACGCGCGTCATGACGAAGAATGCAACACGAGAAACCTGTCGCTCTCTGGGCGCTGA	1140
Db	1086	ACGAGACGCGCGTCATGACGAAGAATGCAACACGAGAAACCTGTCGCTCTCTGGGCGCTGA	1145
QY	1141	TCCTGACACGAGGGCTGTACATTTCTCATGTGAGCAGCTGAGCAAGGGCAACCTGGTGAAC	1200
Db	1146	TCCTGACACGAGGGCTGTACATTTCTCATGTGAGCAGCTGAGCAAGGGCAACCTGGTGAAC	1205
QY	1201	TTCTGCGACCCGGGGGTGAGACCCCTCGTGAACACCGCTCAGCTCCTCAGTTTTCTCTGC	1260
Db	1206	TTCTGCGACCCGGGGGTGAGACCCCTCGTGAACACCGCTCAGCTCCTCAGTTTTCTCTGC	1265
QY	1261	ACGTGCGCGAGGGCATGAGTACTGTGGAGAGCAAGAAAGCTTGTGCACCCGGACTGGCCG	1320
Db	1266	ACGTGCGCGAGGGCATGAGTACTGTGGAGAGCAAGAAAGCTTGTGCACCCGGACTGGCCG	1325
QY	1321	CCCCCAACATCTGTGTCACAGAGACCTGGGGGCAAGTACGAGACTTTGGCCCTGGCCA	1380
Db	1326	CCCCCAACATCTGTGTCACAGAGACCTGGGGGCAAGTACGAGACTTTGGCCCTGGCCA	1385
QY	1381	AAGCCGAGCGGAGGGGCTAGACATCAAGCCGGCTGCCGTCACAGTGAAGCGGCCCGCAGG	1440
Db	1386	AAGCCGAGCGGAGGGGCTAGACATCAAGCCGGCTGCCGTCACAGTGAAGCGGCCCGCAGG	1445
QY	1441	CTCTCAAAACACGGGAAGTTCACACGCAAGTGGATGTCTGGAGTTTGGGGTCTGCTCT	1500
Db	1446	CTCTCAAAACACGGG---GTTCACACGCAAGTGGATGTCTGGAGTTTGGGGTCTGCTCT	1502
QY	1501	GGAGAGCTTCTCATATATGAGACGGGCTCCGCTAAATAATCTACTGAAAGAGGTGTCGG	1560
Db	1503	GGAGAGCTTCTCATATGAGACGGGCTCCGCTAAATAATCTACTGAAAGAGGTGTCGG	1562
QY	1561	AGGCCGTGGAAGAGGGGTACCGCATGGAACCCCGGAGGGCTGTCAAGGCCCGCTGCACG	1620
Db	1563	AGGCCGTGGAAGAGGGGTACCGCATGGAACCCCGGAGGGCTGTCAAGGCCCGCTGCACG	1622
QY	1621	TCCTCATGAGACGTGCTGCTGGAGGACAGAGCCCGCGCCGACACCTTCGCGAAATCG	1680
Db	1623	TCCTCATGAGACGTGCTGCTGGAGGACAGAG--CCGCGCCCGCGCCACCTTCGCGAAATCG	1681
QY	1681	CCGGAAGACTGGCCCGGGAGCTAGCAGTGAAGTGGGCCCAAGCCCTCGCTCAGGGCAGG	1740
Db	1682	CCGGAAGACTGGCCCGGGAGCTAGCAGTGAAGTGGGCCCAAGCCCTCGCTCAGGGCAGG	1741
QY	1741	ACGGCGAGGGCTCCACCTCGGCCCGGAGCCAGGAGGCCCTGACCCACCCGCTGGGGCCCT	1800
Db	1742	ACGGCGAGGG--TCCACCTCGGCCCGGAGGCCAGGAGGCCCTGACCCACCCGCT--GGCCCT	1798
QY	1801	TGGCCCCCAGAGGACCCAGAGAGTGGAGAGTGGCCGCTGGGGGCACTGACCAAGGCCCAAG	1860
Db	1799	TGGCCCCCAGAGGACCCAGAGAGTGGAGAGTGGCCGCTGGGGGCACTGACCAAGGCCCAAG	1858
QY	1861	AGGGTCCAGGGGGGGCAAGTCACTCTCTGGTGGGCCCAAGCAGGGGGCTGGGCCACGTAGGG	1920
Db	1859	AGGGTCCAGGGGGGGCAAGTCACTCTCTGGTGGGCCCAAGCAGGGGGCTGGGCCACGTAGGG	1918
QY	1921	GGCTCTGGGGCGCCGTGGAGACCCCAAGACCTTGGGAAGGATGATCGCCCATAAAGACGG	1980
Db	1919	GGCTCTGGGGCGCCGTGGAGACCCCAAGACCTTGGGAAGGATGATCGCCCATAAAGACGG	1978
QY	1981	ATTCTAAG 1989	
Db	1979	ATTCTAAG 1987	

RESULT 6
US-08-604-989A-11
: Sequence 11, Application US/08604989A
: Patent No. 5834208
: GENERAL INFORMATION:
: APPLICANT: Sakano, S.
: TITLE OF INVENTION: No. 5834208e1 Tyrosine Kinase

NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/604,989A
FILING DATE: February 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Charles E. Miller
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 1920-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1942 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: human
STRAIN: UT-7
US-08-604-989A-11

Query Match 95.28; Score 1903.4; DB 2; Length 1942;
Best Local Similarity 99.78; Pred. No. 0;
Matches 1939; Conservative 0; Mismatches 1; Indels 5; Gaps 3;

52 CGGAGGCGCTCTCTGCGGGGCGGGCGGG--GCGGCTGCGGGGCGCCCTGAGCAGAAAA 109
1 CGGAGGCGCTCTCTGCGGGGCGGGCGGGCGGGCGGGCTGCGGGGCGCCCTGAGCAGAAAA 60
110 CAGGAAGAACAGGCTGCTGTCAGATGTCAGCCAGCTCTCTGTCAGCCAGCCCT 169
61 CAGGAAGAACAGGCTGCTGTCAGATGTCAGCCAGCTCTCTGTCAGCCAGCCACT 120
170 GGGCTGTGCGAGGCGCATTTCCAGCGTCCCGAGCTGTACCACTTGTCTAGTGTCTC 229
121 GGGCTGTGCGAGGCGCATTTCCAGCGTCCCGAGCTGTACCACTTGTCTAGTGTCTC 180
230 ACCGCTCTCAGTTTCCCTCTGCGGGGCGGATGCGGGGCGGAGGCTCTGTTTCTGCGG 289
181 ACCGCTCTCAGTTTCCCTCT-GGGGGCGATGCGGGGCGGAGGCTCTGTTTCTGCGG 239
290 GGCATTTACAGCGGTGTGATTTCTGTAGAACTTCCCGGGGTGAGCCCGCTTCTCCG 349
240 GGCATTTACAGCGGTGTGATTTCTGTAGAACTTCCCGGGGTGAGCCCGCTTCTCCG 299
350 AGCCTGGCACCCCTCTCCGCTCTCAGCAGAGATGCCAAGCAGGCGCTGAGGCCCGGCGAC 409
300 AGCCTGGCACCCCTCTCCGCTCTCAGCAGAGATGCCAAGCAGGCGCTGAGGCCCGGCGAC 359
410 CCGATGTATACCAAAATGCGACAGACACCCGCCCAAGCAGGGAGCTGGCTTTCCGCAA 469
360 CCGATGTATACCAAAATGCGACAGACACCCGCCCAAGCAGGGAGCTGGCTTTCCGCAA 419
470 GGGCGAGTGTGTCACCTCTCTGAGGCGCTCGAAGAACAGAGCTGTACCGCGCTCAAGCA 529
420 GGGCGAGTGTGTCACCTCTCTGAGGCGCTCGAAGAACAGAGCTGTACCGCGCTCAAGCA 479

530 CCACACAGTGTGACAGGAGGGGCTGTGGACGCTGGGGCGCTGCGGAGCGGAGGCCCT 589
480 CCACACAGTGTGACAGGAGGGGCTGTGGACGCTGGGGCGCTGCGGAGCGGAGGCCCT 539
590 CTCGCGAGACCCCAAGCTCAGCCGCTGAGCCGCTGAGGAGAAATCTCGGGCCAGGA 649
540 CTCGCGAGACCCCAAGCTCAGCCGCTGAGCCGCTGAGGAGAAATCTCGGGCCAGGA 599
650 GGGCTGTCCAGACGCTGAGCGCTCCCGAGATGAGGCTGTCTGCTGGGGAGTCCGCGC 709
600 GGGCTGTCCAGACGCTGAGCGCTCCCGAGATGAGGCTGTCTGCTGGGGAGTCCGCGC 659
710 CCACCCCGGAGCTAGCTCTGTGCTGAGCTTTGGCCCGAGCTATCCACTACCGCT 769
660 CCACCCCGGAGCTAGCTCTGTGCTGAGCTTTGGCCCGAGCTATCCACTACCGCT 719
770 GGTGACACCGGAGCGCCACCTCACAATGAGTGAAGCCGCTGTCTGCAACCTCATGGA 829
720 GGTGACACCGGAGCGCCACCTCACAATGAGTGAAGCCGCTGTCTGCAACCTCATGGA 779
830 CATGTGAGACATTTACAGCAAGGACAGGCGCTATCTGACCAAGCTGTGAGACCAA 889
780 CATGTGAGACATTTACAGCAAGGACAGGCGCTATCTGACCAAGCTGTGAGACCAA 839
890 GCGGAAACACGGGACCAAGTGGCCGAGAGAGAGCTGGCCAGCGGCGCTGTATCTGAA 949
840 GCGGAAACACGGGACCAAGTGGCCGAGAGAGAGCTGGCCAGCGGCGCTGTATCTGAA 899
950 CTGACACATTTGACATTTGGAGACAGATTCGAGAGGAGAGTTTGGAGCTGTCCGCA 1009
900 CTGACACATTTGACATTTGGAGACAGATTCGAGAGGAGAGTTTGGAGCTGTCCGCA 959
1010 GGGTGAATACCTGGGGGCAAAAGTGGCGGTGAAGATATCAAGTGTATGTGACAGCCA 1069
960 GGGTGAATACCTGGGGGCAAAAGTGGCGGTGAAGATATCAAGTGTATGTGACAGCCA 1019
1070 GGGCTTCTGAGAGACAGGCGCTCATGACAGAGATCAACACAGCAACTGGTGCCTT 1129
1020 GGGCTTCTGAGAGACAGGCGCTCATGACAGAGATCAACACAGCAACTGGTGCCTT 1079
1130 CTTGGGCGTATCTTGACACAGGCGCTCATGATGTATGATGAGACGTGAGCAAGGCAA 1189
1080 CTTGGGCGTATCTTGACACAGGCGCTCATGATGTATGATGAGACGTGAGCAAGGCAA 1139
1190 CTTGGTGAATTTCTTGCGGACCGGGGTGACAGCCCTTGTAAGACCGCTAGCTCTGCA 1249
1140 CTTGGTGAATTTCTTGCGGACCGGGGTGACAGCCCTTGTAAGACCGCTAGCTCTGCA 1199
1250 GTTTTCTCTGACAGTGGCCGAGGCGATGAGATACCTGGAAGCAAGACTTGTGCACCG 1309
1200 GTTTTCTCTGACAGTGGCCGAGGCGATGAGATACCTGGAAGCAAGACTTGTGCACCG 1259
1310 CGACCTGCGCGCGCGCAACATCTGTCTGAGAGGACCTGTGGGCCAAGGTCAGCGACT 1369
1260 CGACCTGCGCGCGCGCAACATCTGTCTGAGAGGACCTGTGGGCCAAGGTCAGCGACT 1319
1370 TGGCTTGGCCAAAGCCGAGCGGAGAGGGGCTAGACTCAAGCCGCGCTGCCCTCAAGTGGAC 1429
1320 TGGCTTGGCCAAAGCCGAGCGGAGAGGGGCTAGACTCAAGCCGCGCTGCCCTCAAGTGGAC 1379
1430 GGGCGCCGAGGCTCTCAACACGCGGAATTCACACAGCAAGTGGAGTGTGAGATTTGG 1489
1380 GGGCGCCGAGGCTCTCAACACGCGGAATTCACACAGCAAGTGGAGTGTGAGATTTGG 1439
1490 GGTGCTCTCTGGGAGGTCTTCTCATATGAGACGGGCTCCCTACCTTAAATGTCTACCTGAA 1549
1440 GGTGCTCTCTGGGAGGTCTTCTCATATGAGACGGGCTCCCTACCTTAAATGTCTACCTGAA 1499
1550 AGAGGTGTGAGAGCGGTGAGAGAGGGTTCACCATGAGAACCCCGAGGGCTGTCCAGG 1609
1500 AGAGGTGTGAGAGCGGTGAGAGAGGGTTCACCATGAGAACCCCGAGGGCTGTCCAGG 1559
1610 CCCGCTGACAGTCTCTATGAGACAGTGTGAGAGGACAGGCCCGCCCGCCACCTTT 1669

Db 1560 CCCCCTGACGCTCTCATGACACAGCTCTGGAGAGAGCCCGCCGACCTT 1619
QY 1670 CCGCAACTGGCGAGAGAGCTGGCCCGGAGCTACAGAGTGCAGGTGCCCAAGCTCCGT 1729
Db 1620 CCGCAACTGGCGAGAGAGCTGGCCCGGAGCTACAGAGTGCAGGTGCCCAAGCTCCGT 1679
QY 1730 CTCAGGCGAGAGAGCTCCAGCTCCAGCTCCCGGAGAGCTTCAGCCCAACC 1789
Db 1680 CTCAGGCGAGAGAGCTCCAGCTCCAGCTCCCGGAGAGCTTCAGCCCAACC 1739
QY 1790 GGTGGGCGCTTGGCCCGGAGAGAGAGAGTGGAGAGTGGCGCTGGGGGCACTGAC 1849
Db 1740 GGT--GGCCCTTGGCCCGGAGAGAGAGAGTGGAGAGTGGCGCTGGGGGCACTGAC 1797
QY 1850 CAGCCCAAGAGAGAGTCCAGGCGGAGAGTATCTCTGCTGCCACAGAGAGGCTGG 1909
Db 1798 CAGCCCAAGAGAGAGTCCAGGCGGAGAGTATCTCTGCTGCCACAGAGAGGCTGG 1857
QY 1910 CCGACGTAGGGGCTCTGGCGGCGCTGGACACCCAGACCTGCGAAGGATGATCGCC 1969
Db 1858 CCGACGTAGGGGCTCTGGCGGCGCTGGACACCCAGACCTGCGAAGGATGATCGCC 1917
QY 1970 GATTAAGAGGATTTCTAAGACTCT 1994
Db 1918 GATTAAGAGGATTTCTAAGACTCT 1942

RESULT 7
US-08-604-989A-10
Sequence 10, Application US/08604989A
Patent No. 5834208
GENERAL INFORMATION:
APPLICANT: Sakano, S.
TITLE OF INVENTION: NO. 5834208el Tyrosine Kinase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/604,989A
FILING DATE: February 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Charles E. Miller
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 1920-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1521 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: human
STRAIN: UT-7
US-08-604-989A-10
Query Match 76.0%; Score 1519.4; DB 2; Length 1521;

Best Local Similarity 99.9%; Pred. No. 9.5e-300;
Matches 1520; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 258 ATGCGGGGCGAGGCTCTCTGTTCTCTGCGGGCATTTTACGCGCTGTGATCTGCTGAG 317
Db 1 ATGCGGGGCGAGGCTCTCTGTTCTCTGCGGGCATTTTACGCGCTGTGATCTGCTGAG 60
QY 318 GAATTCCTCCGGGGAGCCCCCGCTCTCTCCGAGCTGGGACCCCGCTCCGCTGCAAGC 377
Db 61 GAATTCCTCCGGGGAGCCCCCGCTCTCTCCGAGCTGGGACCCCGCTCCGCTGCAAGC 120
QY 378 AGATGCGCAACGAGGCGCTGGGCGCGGAGCCCGGAGCCAGTGTATACCAATATGCGACAC 437
Db 121 AGATGCGCAACGAGGCGCTGGGCGCGGAGCCCGGAGCCAGTGTATACCAATATGCGACAC 180
QY 438 CGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGCGACGCTGTACCATCTCTGGAGCC 497
Db 181 CGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGCGACGCTGTACCATCTCTGGAGCC 240
QY 498 TGGGAGAAAGAGCTGCTGACCGGCTCAAGCAGCAGCAGTGGAGAGAGGGCTGCTG 557
Db 241 TGGGAGAAAGAGCTGCTGACCGGCTCAAGCAGCAGCAGTGGAGAGAGGGCTGCTG 300
QY 558 GCAGCTGGGGGCTGCGGGAGCGGGAGGAGCCCTCTCCGAGACCCCAAGCTCAGCCTCATG 617
Db 301 GCAGCTGGGGGCTGCGGGAGCGGGAGGAGCCCTCTCCGAGACCCCAAGCTCAGCCTCATG 360
QY 618 CCGTGTCTTCCAGGGGAAGATCTCGGGCCAGAGAGCTGTCCAGCACTGCACTCCCGAG 677
Db 361 CCGTGTCTTCCAGGGGAAGATCTCGGGCCAGAGAGCTGTCCAGCACTGCACTCCCGAG 420
QY 678 GATGGGCTGTCGTTGGTGGGGAGTCCGGCGGCAACCCCGGAGCTACGTCCTGGGCTG 737
Db 421 GATGGGCTGTCGTTGGTGGGGAGTCCGGCGGCAACCCCGGAGCTACGTCCTGGGCTG 480
QY 738 AGCTTTGGCCGCGAGCTATCCACTACCGCTGTGCGACGCGAGCCCACTCATCATC 797
Db 481 AGCTTTGGCCGCGAGCTATCCACTACCGCTGTGCGACGCGAGCCCACTCATCATC 540
QY 798 GATGAGGCGCTGTTCTTCTGCAACCTCATGAGCATGTGTGAGCATTTACAGCAGACAG 857
Db 541 GATGAGGCGCTGTTCTTCTGCAACCTCATGAGCATGTGTGAGCATTTACAGCAGACAG 600
QY 858 GGGCTATCTGACCAAGCTGTGTAGACCAAGGGAACCGGGACCAAGTCCGGAG 917
Db 601 GGGCTATCTGACCAAGCTGTGTAGACCAAGGGAACCGGGACCAAGTCCGGAG 660
QY 918 GAGAGCTGGCCAGGGGGGCTGTTACTGAACTGCAAGCTTGTGACATTTGGAGCAGAG 977
Db 661 GAGAGCTGGCCAGGGGGGCTGTTACTGAACTGCAAGCTTGTGACATTTGGAGCAGAG 720
QY 978 ATCGGAGAGGAGAGTGTGAGCTGTCTGCAAGGCTGATACCTTGGGGCAAAAGGTGCC 1037
Db 721 ATCGGAGAGGAGAGTGTGAGCTGTCTGCAAGGCTGATACCTTGGGGCAAAAGGTGCC 780
QY 1038 GTGAAGATATCAAGTGTGATGTGACAGCCAGGCTTCTGAGAGAGAGGCGCTCATG 1097
Db 781 GTGAAGATATCAAGTGTGATGTGACAGCCAGGCTTCTGAGAGAGAGGCGCTCATG 840
QY 1098 ACGAAGATGCAACAGAGAACTGTGCTCTGCGGCGTGTGATCTGCAACAGAGGCTG 1157
Db 841 ACGAAGATGCAACAGAGAACTGTGCTCTGCGGCGTGTGATCTGCAACAGAGGCTG 900
QY 1158 TACATTTGTATGAGCAGCTGAGCAAGGGCAACTGTGTAATTTGCGGAGCCGGGGT 1217
Db 901 TACATTTGTATGAGCAGCTGAGCAAGGGCAACTGTGTAATTTGCGGAGCCGGGGT 960
QY 1218 CGAGCCCTGCTGAGACCGGCTCAGCTCTGAGTTTCTGCAAGTGGCGGAGGGATG 1277
Db 961 CGAGCCCTGCTGAGACCGGCTCAGCTCTGAGTTTCTGCAAGTGGCGGAGGGATG 1020
QY 1278 GAGTACCTGAGAGCAAGAGCTGTGACCGGCAAGCTGGCGGCGCAACATCTGCTC 1337
Db 1337 GAGTACCTGAGAGCAAGAGCTGTGACCGGCAAGCTGGCGGCGCAACATCTGCTC

Db 1021 GAGTACTGAGACGAAGAGCTTTGTGACCGGACCTGGCCGCCGCAACATCTGGTC 1080
QY 1338 TCAGAGGACCTGTGTGGCCAAAGTCAAGGACTTTGGCTGGCCAAAGCCGAGGAGGG 1397
Db 1081 TCAGAGGACCTGTGTGGCCAAAGTCAAGGACTTTGGCTGGCCAAAGCCGAGGAGGG 1140
QY 1398 CTGAGCTCAAGCCGGCTGCCCTGCAAGTGGAGCGGCCGAGGCTCTCAAAACGAGGAG 1457
Db 1141 CTGAGCTCAAGCCGGCTGCCCTGCAAGTGGAGCGGCCGAGGCTCTCAAAACGAGGAG 1200
QY 1458 TTCACGACGAGTGTGATGTCTGAGTCTTTGGGGTGTGCTGGGGAGGCTCTCATAT 1517
Db 1201 TTCACGACGAGTGTGATGTCTGAGTCTTTGGGGTGTGCTGGGGAGGCTCTCATAT 1260
QY 1518 GAGAGGGGCTCCGTACCTAAATGTCACTGAAAGAGGTGTGAGAGCCGTGAGAGAGGG 1577
Db 1261 GAGAGGGGCTCCGTACCTAAATGTCACTGAAAGAGGTGTGAGAGCCGTGAGAGAGGG 1320
QY 1578 TACCGATGGAACCCCCCGAGGGCTGTCCAGGCCCCGTGACGCTCTCATGAGCAGCTGC 1637
Db 1321 TACCGATGGAACCCCCCGAGGGCTGTCCAGGCCCCGTGACGCTCTCATGAGCAGCTGC 1380
QY 1638 TGGGAGGACAGAGCCCGCCGCGCCGACCCCTTCGCAAACTGGCCGAGAAAGTGGGCCGG 1697
Db 1381 TGGGAGGACAGAGCCCGCCGCGCCGACCCCTTCGCAAACTGGCCGAGAAAGTGGGCCGG 1440
QY 1698 GAGCTACGAGAGTGTGAGTGGCCGAGGCTCGCTCTAGGGGAGGAGCGCCGAGGCTTCACC 1757
Db 1441 GAGCTACGAGAGTGTGAGTGGCCGAGGCTCGCTCTAGGGGAGGAGCGCCGAGGCTTCACC 1500
QY 1758 TCGCCCCGAAAGCCAGAGGCC 1778
Db 1501 TCGCCCCGAAAGCCAGAGGCC 1521

RESULT 8
US-08-604-989A-9
Sequence 9, Application US/08604989A
Patent No. 5834208
GENERAL INFORMATION:
APPLICANT: Sakano, S.
TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/604,989A
FILING DATE: February 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Charles E. Miller
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 1920-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141.PENNIE
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1398 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: human
STRAIN: UT-7
US-08-604-989A-9

Query Match 69.9%; Score 1398; DB 2; Length 1398;
Best Local Similarity 100.0%; Pred. No. 4e-275;
Matches 1398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 ATGCCAAGAGAGCGCGGCGCCCGGCGCCGACCCAGTATATACCAAAAGCGAGACACCCCG 440
Db 1 ATGCCAAGAGAGCGCGGCGCCCGGCGCCGACCCAGTATATACCAAAAGCGAGACACCCCG 60
QY 441 CCCAAGCCAGAGGAGCGTGCCTTCCCAAGAGGCGAGCGTGTACATCTCTGAGGCGTGC 500
Db 61 CCCAAGCCAGAGGAGCGTGCCTTCCCAAGAGGCGAGCGTGTACATCTCTGAGGCGTGC 120
QY 501 GAGACAAAGAGCTGTATCCGCGTCAAGACACACACACAGTGTGAGACAGAGGCGTGTGCA 560
Db 121 GAGACAAAGAGCTGTATCCGCGTCAAGACACACACAGTGTGAGACAGAGGCGTGTGCA 180
QY 561 GCTGGGGGCGCTGGGGAGGGGGAGGGCGCTCTCCGAGACCCCAAGCTCAGCCCTCATGGCG 620
Db 181 GCTGGGGGCGCTGGGGAGGGGGAGGGCGCTCTCCGAGACCCCAAGCTCAGCCCTCATGGCG 240
QY 621 TGGTTCCAGGGGAGATCTCGGGCCAGAGAGGTGTTCAGCAGCTGCAGGCTCCCGAGAGAT 680
Db 241 TGGTTCCAGGGGAGATCTCGGGCCAGAGAGGTGTTCAGCAGCTGCAGGCTCCCGAGAGAT 300
QY 681 GGGCTGTCTCTGTGTCGGGAGTCCGGCGCCACCCCGGAGCTAGCTCTGTGCTGTGAGC 740
Db 301 GGGCTGTCTCTGTGTCGGGAGTCCGGCGCCACCCCGGAGCTAGCTCTGTGCTGTGAGC 360
QY 741 TTTGGCCGGGAGGTCATCCACTACCGGCTGTGTCAGCAGGAGGCGCCACTCCAAATCAT 800
Db 361 TTTGGCCGGGAGGTCATCCACTACCGGCTGTGTCAGCAGGAGGCGCCACTCCAAATCAT 420
QY 801 GAGGCGGCTTCTTCTGCAACCTCATGAGCATGTGTGAGCATTACAGCAAGAGAGGCG 860
Db 421 GAGGCGGCTTCTTCTGCAACCTCATGAGCATGTGTGAGCATTACAGCAAGAGAGGCG 480
QY 861 GCTATCTGCACCAAGCTGTGTGAGACCAAGGCGGAAACAGGAGCAAGTCCGCGAGAG 920
Db 481 GCTATCTGCACCAAGCTGTGTGAGACCAAGGCGGAAACAGGAGCAAGTCCGCGAGAG 540
QY 921 GAGCTGGCCAGGGGCGGGCTGTGAGACCTGTGAGACCTGTGAGACCTGTGAGACCTGTG 980
Db 541 GAGCTGGCCAGGGGCGGGCTGTGAGACCTGTGAGACCTGTGAGACCTGTGAGACCTGTG 600
QY 981 GGAAGAGGAGAGTTTGGAGCTGTCTGCAAGGCTGTGAGTACCTGGGCAAAAGGTGGCCGTG 1040
Db 601 GGAAGAGGAGAGTTTGGAGCTGTCTGCAAGGCTGTGAGTACCTGGGCAAAAGGTGGCCGTG 660
QY 1041 AAGATATCAAGTGTATGTGACAGCCAGGCTTCTGTGAGAGAGGCGCTCATGACG 1100
Db 661 AAGATATCAAGTGTATGTGACAGCCAGGCTTCTGTGAGAGAGGCGCTCATGACG 720
QY 1101 AAGATGCAACAGAGAACTGTGTGCTGCGGGGGGTGATCTGACCAAGGCGGTGATC 1160
Db 721 AAGATGCAACAGAGAACTGTGTGCTGCGGGGGGTGATCTGACCAAGGCGGTGATC 780
QY 1161 ATTGTATGAGACAGTGTGAGCAAGGCAACCTGTGTAATTTCTGCGAGCCCGGGGTGCA 1220
Db 781 ATTGTATGAGACAGTGTGAGCAAGGCAACCTGTGTAATTTCTGCGAGCCCGGGGTGCA 840
QY 1221 GCCCTGTGTAACACGCGTCAAGTCTGTGCAAGTTTCTGTGACAGTGTGAGCGAGATGAG 1280
Db 841 GCCCTGTGTAACACGCGTCAAGTCTGTGCAAGTTTCTGTGACAGTGTGAGCGAGATGAG 900
QY 1281 TACCTGGAGAGCAAGAGCTGTGACCGCGAGCTGGCCGCCGCAACATCTGTGCTCA 1340
Db 901 TACCTGGAGAGCAAGAGCTGTGACCGCGAGCTGGCCGCCGCAACATCTGTGCTCA 960

QY	1341	GAGGACCTGGTGGCCCAAGCTACGACATTGGCTGGCCAAAGCCGAGCGGAAGGGGCTA	1400
Db	961	GAGGACCTGGTGGCCCAAGCTACGACATTGGCTGGCCAAAGCCGAGCGGAAGGGGCTA	1020
QY	1401	GACTCAAGCCGCGCTGCCCTCACTGAGCGGCGCCAGAGGCTCTCAACACAGGGAATTC	1460
Db	1021	GACTCAAGCCGCGCTGCCCTCACTGAGCGGCGCCAGAGGCTCTCAACACAGGGAATTC	1080
QY	1461	ACCAAGCAAGTCGGATGTCGTGGAGTTTGGGTGCTGCTCTGGGAAGTCTTCTCATATGGA	1520
Db	1081	ACCAAGCAAGTCGGATGTCGTGGAGTTTGGGTGCTGCTCTGGGAAGTCTTCTCATATGGA	1140
QY	1521	CGGAGCTCCGTACCTTAAATATGTCACTGAAAGAGTGTGAGAGGCCGTGGAGAAAGGGTAC	1580
Db	1141	CGGAGCTCCGTACCTTAAATATGTCACTGAAAGAGTGTGAGAGGCCGTGGAGAAAGGGTAC	1200
QY	1581	CGCATGGAAACCCCGCAGAGGCTGTCCAGAGGCCCGTGCACGTCTCATAGACAGCTGCTGG	1640
Db	1201	CGCATGGAAACCCCGCAGAGGCTGTCCAGAGGCCCGTGCACGTCTCATAGACAGCTGCTGG	1260
QY	1641	GAGGCAGAGCCCGCGCGCGGCCACCCCTCCGGCAATGGCGCGGAGAGCTGGCCCGGGAG	1700
Db	1261	GAGGCAGAGCCCGCGCGCGGCCACCCCTCCGGCAATGGCGCGGAGAGCTGGCCCGGGAG	1320
QY	1701	CTACGAGATGAGGATGGCCCAAGCCTTCGCTCTCAAGGAGAGACAGCAGCGCTCCACCTCG	1760
Db	1321	CTACGAGATGAGGATGGCCCAAGCCTTCGCTCTCAAGGAGAGACAGCAGCGCTCCACCTCG	1380
QY	1761	CCCCGAAGCCACGAGAGCCC	1778
Db	1381	CCCCGAAGCCACGAGAGCCC	1398

```

RESULT 9
US-09-741-154-1
: Sequence 1, Application US/09741154
: Patent No. 6437110
: GENERAL INFORMATION:
: APPLICANT: BEASLEY, Ellen M. et al
: TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
: TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
: TITLE OF INVENTION: THEREOF
: FILE REFERENCE: CLO01061
: CURRENT APPLICATION NUMBER: US/09/741,154
: CURRENT FILING DATE: 2000-12-21
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 1713
: TYPE: DNA
: ORGANISM: Human,
US-09-741-154-1

```

Query Match	68.8%	Score 1377	DB 4	Length 1713
Best Local Similarity	99.38%	Pred. No. 7	6e-271	
Matches 1383; Conservative	0	Mismatches	10	Indels 0; Gaps 0;

QY	608	CAGCCTCATGCGCTGTTCACAGGGAGATCTCCGGCCAGGAGGCTGTCCAGCACTGC	667
Db	297	CAGCAGCTTCTTGTTGCCACGGGAAAGATCTCGGGCAGAGGACTGTCCAGCACTGC	356
QY	668	GCCCTCCGAGGATGGGCTGTTCCTGTGTGCGGGAATCCGGCGCCACC GGCGCACTACGT	727
Db	357	GCCTCCGAGGATGGGCTGTTCCTGTGTGCGGGAATCCGGCGCCACC GGCGCACTACGT	416
QY	728	CCTGTGCGTAGCTTTTGAGCGCGAAGCATCATTAACGGCTGTGTGCACCGCGAGGGA	787
Db	417	CCTGTGCGTAGCTTTTGAGCGCGAAGCATCATTAACGGCTGTGTGCACCGCGAGGGA	476
QY	788	CCTCACATGAGTAGGCGCGTGTCTTCTGTCAACCTCATGAGCATGGTGTGAGCATTTAC	847
Db	477	CCTCACATGAGTAGGCGCGTGTCTTCTGTCAACCTCATGAGCATGGTGTGAGCATTTAC	536

QY	848	CAAGGACAAGGGCCCTATCTGACCAACAGCTGGTGAACCAAGCGGAAACACGGGACC	907
Dp	537	CAAGGACAAGGGCCCTATCTGACCAACAGCTGGTGAACCAAGCGGAAACACGGGACC	596
QY	908	GTGCGCCGAGAGAGAGCTGGGCGAGGGGGGGCTGTACTGAACCTGGAGCATTTGACATT	967
Dp	597	GTGCGCCGAGAGAGAGCTGGGCGAGGGGGGGCTGTACTGAACCTGGAGCATTTGACATT	656
QY	968	GGAGGACACAGATCGGAGAGGAGGAGTTTGAGAGCTGTCTGTGAGGGGTAGTACCTGGGGCA	1027
Dp	657	GGAGGACACAGATCGGAGAGGAGGAGTTTGAGAGCTGTCTGTGAGGGGTAGTACCTGGGGCA	716
QY	1028	AAAGGTGGCCGTGAAGAAATATCAAGTGTGATGTACAGAGCCAGGCCCTTCTGGAGCGAGAC	1087
Dp	717	AAAGGTGGCCGTGAAGAAATATCAAGTGTGATGTACAGAGCCAGGCCCTTCTGGAGCGAGAC	776
QY	1088	GGCCGTCATGACGAAGATGCAACACGAGAACCTTGATGAGCTCTCTGGGCGGTGATCTGCA	1147
Dp	777	GGCCGTCATGACGAAGATGCAACACGAGAACCTTGATGAGCTCTCTGGGCGGTGATCTGCA	836
QY	1148	CCAGGGGCTGTACATTGTCATGTGAGACAGTGTGAGCAAGGGCAACTGGTGAACCTTTCGTGCG	1207
Dp	837	CCAGGGGCTGTACATTGTCATGTGAGACAGTGTGAGCAAGGGCAACTGGTGAACCTTTCGTGCG	896
QY	1208	GACCCGGGGGTGACAGCCCTCGTGAACACACCGCTCAGCTCTGTGAGTTTCTGTGACAGTGGC	1267
Dp	897	GACCCGGGGGTGACAGCCCTCGTGAACACACCGCTCAGCTCTGTGAGTTTCTGTGACAGTGGC	956
QY	1268	CGAGGCGATGAGGATACCTGTGAGAGCAGACAAAGACTTGTGCACCCGCGACCTGGCCCGCCAA	1327
Dp	957	CGAGGCGATGAGGATACCTGTGAGAGCAGACAAAGACTTGTGCACCCGCGACCTGGCCCGCCAA	1016
QY	1328	CATCTGTGTGTCAAGAGACCTGTGGGGCAAGGATGAGGAGCATTTGGGCGGCAAGGCGCA	1387
Dp	1017	CATCTGTGTGTCAAGAGACCTGTGGGGCAAGGATGAGGAGCATTTGGGCGGCAAGGCGCA	1076
QY	1388	GCGGAGAGGGGCTAGACTCAAGCCCGGCTGCCCGTCAAGTGTGACGGGCGCCGAGGCTCTCAA	1447
Dp	1077	GCGGAGAGGGGCTAGACTCAAGCCCGGCTGCCCGTCAAGTGTGACGGGCGCCGAGGCTCTCAA	1136
QY	1448	ACAGGGGAAGTTCACACAGCAAGTGGGATGTGAGTGTGGGGTGGCTGCTGTGGGAGGT	1507
Dp	1137	ACAGGGGAAGTTCACACAGCAAGTGGGATGTGAGTGTGGGGTGGCTGCTGTGGGAGGT	1196
QY	1508	CTTCTCATATGAGACGGGCTCCGTACCCCTAAATATCTACTGAAAGAGTGTGTGAGGGCGT	1567
Dp	1197	CTTCTCATATGAGACGGGCTCCGTACCCCTAAATATCTACTGAAAGAGTGTGTGAGGGCGT	1256
QY	1568	GGAGAGAGGGGTACCGCATGGAAACCCCGAGGGCTGTTCAGAGCCCGCTGCACGTCTCAT	1627
Dp	1257	GGAGAGAGGGGTACCGCATGGAAACCCCGAGGGCTGTTCAGAGCCCGCTGCACGTCTCAT	1316
QY	1628	GAGAGAGTGTGGGAGGAGAGAGCCCGGCGCGGCGCACACCTTTCGGCAAACTGGCGGAGAA	1687
Dp	1317	GAGAGAGTGTGGGAGGAGAGAGCCCGGCGCGGCGCACACCTTTCGGCAAACTGGCGGAGAA	1376
QY	1688	GCTTGGCCCGGAGAGTACGCAATGCAAGTGGCCCAAGCTTCGCTTCAGAGGCGAGAGCGCA	1747
Dp	1377	GCTTGGCCCGGAGAGTACGCAATGCAAGTGGCCCAAGCTTCGCTTCAGAGGCGAGAGCGCA	1436
QY	1748	CGGCTTCACCTCGGCCCCGGAAGCCAGAGAGCCCTGACACCCACCCGGTGGGGGCCCTTGGCCCC	1807
Dp	1437	CGGCTTCACCTCGGCCCCGGAAGCCAGAGAGCCCTGACACCCACCCGGTGGGGGCCCTTGGCCCC	1496
QY	1808	AGAGGACCGAGAGAGTGGAGAGTGGCGGCTGTGGGGGCACTGTACACAGGGCCCAAGAGAGGGTTC	1867
Dp	1497	AGAGGACCGAGAGAGTGGAGAGTGGCGGCTGTGGGGGCACTGTACACAGGGCCCAAGAGAGGGTTC	1556
QY	1868	AGGCGGGCAAGTACTCTCTGCTGGTGGCCACAGCAGGGGGCTGGCCACGTAGAGGGGCTCTG	1927
Dp	1557	AGGCGGGCAAGTACTCTCTGCTGGTGGCCACAGCAGGGGGCTGGCCACGTAGAGGGGCTCTG	1616

QY 1928 GCGGCGCCGTGGACACCCGACACTGCGAAGATGATCCGCCATAAAGACGATTCTAA 1987
Db 1617 GCGGCGCCGTGGACACCCGACACTGCGAAGATGATCCGCCATAAAGACGATTCTAA 1676
QY 1988 GGACTTAATAA 2000
Db 1677 GGAATAAATAA 1689

RESULT 10
US-08-604-989A-8
Sequence 8, Application US/08604989A
Patent No. 5834208
GENERAL INFORMATION:
APPLICANT: Sakano, S.
TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/604,989A
FILING DATE: February 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Charles E. Miller
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 1920-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 738 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: human
STRAIN: UT-7
US-08-604-989A-8

Query Match 36.9%; Score 738; DB 2; Length 738;
Best Local Similarity 100.0%; Pred. No. 3e-141;
Matches 738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 954 CACGATTGACATTGGGAGCAGATCGAGAGGAGAGTGTGAGCTGCTGACAGGT 1013
Db 1 CACGATTGACATTGGGAGCAGATCGAGAGGAGAGTGTGAGCTGCTGACAGGT 60
QY 1014 GAGTACCTGGGGCAAAAGTGGCGGTGAGAAATATCAAGTGTGATGACAGGCCAGGCC 1073
Db 61 GAGTACCTGGGGCAAAAGTGGCGGTGAGAAATATCAAGTGTGATGACAGGCCAGGCC 120
QY 1074 TTCTGGAGAGAGCGCGCTCATGAGAGAGATCAACACAGAACTGGTGCCTCTCTG 1133
Db 121 TTCTGGAGAGAGCGCGCTCATGAGAGAGATCAACACAGAACTGGTGCCTCTCTG 180
QY 1134 GCGGTATCTGACACGAGGGGTGTACATGTGATGAGAGCACTGAGCAAGGCAACCTG 1193
Db 181 GCGGTATCTGACACGAGGGGTGTACATGTGATGAGAGCACTGAGCAAGGCAACCTG 240

QY 1194 GTGAACTTTTGGGGAGCCCGGGGTGAGACCCCTGCTGAACACCGCTCAGCTTGCAGTTT 1253
Db 241 GTGAACTTTTGGGGAGCCCGGGGTGAGACCCCTGCTGAACACCGCTCAGCTTGCAGTTT 300
QY 1254 TCTGTGACGTGGCGGAGGAGCATGAGTACTGAGAGCAAGAAAGTGTGACACCGGAG 1313
Db 301 TCTGTGACGTGGCGGAGGAGCATGAGTACTGAGAGCAAGAAAGTGTGACACCGGAG 360
QY 1314 CTGGCCGCGGCAACATCTGTCTCAGAGAGCTGGTGGCCAGAGTCAAGCTTGTGGC 1373
Db 361 CTGGCCGCGGCAACATCTGTCTCAGAGAGCTGGTGGCCAGAGTCAAGCTTGTGGC 420
QY 1374 CTGGCCAAAGCCGAGCGGAAGGGGCTAGACTCAAGCCGCTGCCCTCAAGTGGAGCGG 1433
Db 421 CTGGCCAAAGCCGAGCGGAAGGGGCTAGACTCAAGCCGCTGCCCTCAAGTGGAGCGG 480
QY 1434 CCGGAGGCTCTCAAAACAGGGGAAGTTCACAGCAAGTGGATGTGGAATTTGGGGTG 1493
Db 481 CCGGAGGCTCTCAAAACAGGGGAAGTTCACAGCAAGTGGATGTGGAATTTGGGGTG 540
QY 1494 CTGCTGGGAGGCTCTCATATGAGACGGGCTCCGTAACCTTAAATGTCACTGAAGAG 1553
Db 541 CTGCTGGGAGGCTCTCATATGAGACGGGCTCCGTAACCTTAAATGTCACTGAAGAG 600
QY 1554 GTGTGGAGGCGGTGGAAGGGGTATCCGATGAAACCCCGAGGGCTGTCCAGGCCCC 1613
Db 601 GTGTGGAGGCGGTGGAAGGGGTATCCGATGAAACCCCGAGGGCTGTCCAGGCCCC 660
QY 1614 GTGCAGCTCTCATGAGCAGTCTGTGGAGGAGAGCCCGCGCGGCAACCTTCCG 1673
Db 661 GTGCAGCTCTCATGAGCAGTCTGTGGAGGAGAGCCCGCGCGGCAACCTTCCG 720
QY 1674 AAATGCGCGAGAAAGCTG 1691
Db 721 AAATGCGCGAGAAAGCTG 738

RESULT 11
US-09-741-154-3
Sequence 3, Application US/09741154
Patent No. 6437110
GENERAL INFORMATION:
APPLICANT: BEASLEY, Ellen M. et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: C1001061
CURRENT APPLICATION NUMBER: US/09/741,154
CURRENT FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 16389
TYPE: DNA
ORGANISM: Human
US-09-741-154-3

Query Match 22.8%; Score 455; DB 4; Length 16389;
Best Local Similarity 100.0%; Pred. No. 1.3e-83;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1541 GTCACTGAAAGAGGTGTGAGGCGGTGAGAAAGGGGTACCGCATGAAACCCCGAGGG 1600
Db 13941 GTCACTGAAAGAGGTGTGAGGCGGTGAGAAAGGGGTACCGCATGAAACCCCGAGGG 14000
QY 1601 CTGTCCAGGCGCCGCTGACAGTCTCTATGAGACACTCTGTGGAGGAGAGCCCGCGCG 1660
Db 14001 CTGTCCAGGCGCCGCTGACAGTCTCTATGAGACACTCTGTGGAGGAGAGCCCGCGCG 14060
QY 1661 GCCACCTTCCGAAACTGGCCGAGAAAGTGGCCGAGAAAGTGGCCGAGAAAGTGGCCGAG 1720
Db 14061 GCCACCTTCCGAAACTGGCCGAGAAAGTGGCCGAGAAAGTGGCCGAGAAAGTGGCCGAG 14120

QY 1721 AGCCTCGCTCAGGAGGAGACGCCAGCGCTCCACCTCGCCCCGAAGCCAGAGCCCTG 1780
Db 14121 AGCCTCGCTCAGGAGGAGACGCCAGCGCTCCACCTCGCCCCGAAGCCAGAGCCCTG 14180
QY 1781 ACCCCACCCGGTGGGGCCCTTGGCCCCAGAGAGAGAGAGTGGAGAGTGGCGGTGGG 1840
Db 14181 ACCCCACCCGGTGGGGCCCTTGGCCCCAGAGAGAGAGTGGAGAGTGGCGGTGGG 14240
QY 1841 GGCACCTAGCAGGCGCCAGAGAGAGGTCCAGGGGGGAGTCCATCTCCTGGTGGCCACAGC 1900
Db 14241 GGCACCTAGCAGGCGCCAGAGAGAGGTCCAGGGGGGAGTCCATCTCCTGGTGGCCACAGC 14300
QY 1901 AGGGGCTGGGCCACAGTGAAGGGGCTCTGGGGGCCCGTGGACACCCAGACCTGGCAGGA 1960
Db 14301 AGGGGCTGGGCCACAGTGAAGGGGCTCTGGGGGCCCGTGGACACCCAGACCTGGCAGGA 14360
QY 1961 TGATCGCCCGATTAAGACGATTTCTAAGGACTCTA 1995
Db 14361 TGATCGCCCGATTAAGACGATTTCTAAGGACTCTA 14395

RESULT 12

US-08-604-989A-7
Sequence 7, Application US/08604989A
Patent No. 5834208
GENERAL INFORMATION:
APPLICANT: Sakano, S.
TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/604,989A
FILING DATE: February 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Charles E. Miller
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 1920-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 225 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: human
STRAIN: UT-7
US-08-604-989A-7

Query Match 11.2%; Score 225; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 3e-37;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 621 TGGTTCCACGGGAGAGATCTCGGGCCAGAGGCTGTCCAGCAGCTGCAGCTCCCGGAGAT 680
Db 1 TGGTTCCACGGGAGAGATCTCGGGCCAGAGGCTGTCCAGCAGCTGCAGCTCCCGGAGAT 60

QY 681 GGGCTGTTCCTGATGCGGGAGATCCGCGCCACCACCCGGCGACTGCTGTGGGTAGC 740
Db 61 GGGCTGTTCCTGATGCGGGAGATCCGCGCCACCACCCGGCGACTGCTGTGGGTAGC 120
QY 741 TTTGGCCGCGAGCTGATTCACATACCGCGTGTGTCGACCGCGAGCGCCACCTTCACATGAT 800
Db 121 TTTGGCCGCGAGCTGATTCACATACCGCGTGTGTCGACCGCGAGCGCCACCTTCACATGAT 180
QY 801 GAGGCGGTGTCTTCCTGCAACCTTCATGACATGTTGGAGGACTTAC 845
Db 181 GAGGCGGTGTCTTCCTGCAACCTTCATGACATGTTGGAGGACTTAC 225

RESULT 13

US-07-820-011A-3
Sequence 3, Application US/07820011A
Patent No. 533615
GENERAL INFORMATION:
APPLICANT: Bell, Leonard
APPLICANT: Madri, Joseph A.
APPLICANT: Warren, Stephen L.
APPLICANT: Luthringer, Daniel J.
TITLE OF INVENTION: Genetically Engineered
TITLE OF INVENTION: Endothelial Cells Exhibiting Enhanced
TITLE OF INVENTION: Migration
TITLE OF INVENTION: and Plasmidogen Activator Activity
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 Inch, 360 KB storage
COMPUTER: IBM PC XT
OPERATING SYSTEM: PC-DOS/MS-DOS 2.10
SOFTWARE: Displaywrite 3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/820,011A
FILING DATE: 19920106
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: LB-101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255 1400
TELEFAX: (203) 254 1101
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1611
TYPE: NUCLEIC ACID
STRANDEDNESS: Double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapien
POSITION IN GENOME:
CHROMOSOME/SEGMENT: Chromosome 20
PUBLICATION INFORMATION:
AUTHORS: Anderson, Stephen K.
AUTHORS: Gibbs, Carol P.
AUTHORS: Tanaka, Akio
AUTHORS: Kung, Hsing-Jien
AUTHORS: Fujita, Donald J.
TITLE: Human Cellular Src Gene:
TITLE: Nucleotide Sequence and Derived Amino
TITLE: Acid Sequence of the Region Coding for
TITLE: the Carboxy-Terminal Two-Thirds of

TITLE: pp60c-src
JOURNAL: Molecular and Cellular Biology
VOLUME: 5
ISSUE: 5
PAGES: 1122-1129
DATE: May, 1985
PUBLICATION INFORMATION:
AUTHORS: Tanaka, Akio
AUTHORS: Gibbs, Carol P.
AUTHORS: Arthur, Richard R.
AUTHORS: Anderson, Stephen K.
AUTHORS: Kung, Hsing-Jien
Fujita, Donald J.
TITLE: DNA Sequence Encoding the
TITLE: Amino-Terminal Region of the Human c-src
TITLE: Protein: Implications of Sequence
TITLE: Divergence among src-Type Kinase
TITLE: Oncogenes
JOURNAL: Molecular and Cellular Biology
VOLUME: 7
ISSUE: 5
PAGES: 1978-1983
DATE: May, 1987
US-07-820-011A-3

Query Match 10.6%; Score 212.4; DB 1; Length 1611;
Best Local Similarity 55.9%; Pred. No. 1.5e-34;
Matches 455; Conservative 0; Mismatches 341; Indels 18; Gaps 2;

QY 897 CACGGACCAAGTCGGCCGAGAGAGCTGGCCGAGGCGGGCTGTTACTGAACCTGCAG 956
DB 745 CCCACGTCGAAGCCGAGACTCAGGGCTGGCCAAAGGATCCCTGGAGATCCCTCGGAG 804
QY 957 CATTGTACATTTGGAGACAGATCGAGAGAGAGAGTTTGAGACTTCTCGAGGGTAG 1016
DB 805 TGGCTGGGCTGGAGTCAAGCTGGCCAGGGCTGTTGGCAGAGTGGATGGGAGC 864
QY 1017 TACCTGGGCAAAAGGTGGCCCTGAAGAAATATCAAGTGTGATG-----TGACAGCCAG 1070
DB 865 TGAACGGTACCAACAGGGGTGGCCATCAAAACCTGAACCTGGACAGATGTTCCAGAG 924
QY 1071 GCCTTCCTGGAGAGAGGCGCCCTCATAGAGAGATGCAACAGAGAACTGGTGCCTGC 1130
DB 925 GCCTTCCTGGAGAGGCGCCAGTCAAGAAAGCTGAGAGATGAGAAAGCTGGCAGTTG 984
QY 1131 CTGGGGGTATCTCTGACCAAGGGGCTGTATGTTGATGAGACAGCTGAGCAAGGCAAC 1190
DB 985 TATGCTGGTTCAGAGAGAGGCCATTTATGTCAGAGATGATGAGCAAGGGAGT 1044
QY 1191 CTGGTGAATTTCTGCGGAGCCGGGGTGAAGCCCTGTAACACCGCTAGCTCTGCAG 1250
DB 1045 TTGCTGGACTTCTCAAGGGGAGACAGCAAGTACCTCGGCTGCTCAGCTGGTGGAC 1104
QY 1251 TTTTCTCTGACGTGGAGGAGGATGATGATCCTGAGAGAGAAAGTGTGGACCGGC 1310
DB 1105 ATGGCTGCTCAATGCTCTCAAGGATGAGTGGAGAGGATGATGATGATGATGATGAT 1164
QY 1311 GACCTGGGCGCCGCAATCTCTGATGAGAGAGAGTGGTGGCAAGGTCAGCACTTT 1370
DB 1165 GACCTTCTGAGCAAGATCTCTGATGAGAGAGAGTGGTGGCAAGGTCAGCACTTT 1224
QY 1371 GGCCTGGCCAAAGCC-----GAGCGAAGGGGCTGACTCAAGCCGGCTGCC 1418
DB 1225 GGGCTGGCTGGCTCTTGAAGACATGATGATGATGATGATGATGATGATGATGATGAT 1284
QY 1419 GTCAAGTGAAGGCGCGGAGGCTCAAAACAGGGAATTTACCAAGCAAGTGCAGTGC 1478
DB 1285 ATCAAGTGAAGGCTCAAGAGCTGCTCTATGCGCGCTTCAAGTCAAGTGGAGTG 1344
QY 1479 TGAAGTGGGCTGCTCTGAGAGTCTTCTATGAGAGAGGCTCCGTAACCTTAA 1538
DB 1345 TGGTCTTGGGATCTGCTGACTGATGATGATGATGATGATGATGATGATGATGATGAT 1404

QY 1539 ATGTCACTGAAGAGGTGTGAGAGCCGTGGAGAAAGGGTACCAGCATGGAACCCCGAG 1598
DB 1405 ATGTCAAGCCGAGGCTGTGAGCAGTGGAGAGCGGGCTACCGGATGCCCTGCCCGCG 1464
QY 1599 GCGTGTCAAGGCCCTGTGACATCTCTCAAGAGAGTGTGGAGAGCAAGCCCGCGCC 1658
DB 1465 GAGTGTCCGAGTCTCTGACAGCACTCATGATGCTGCTGCGGAAAGAGCTGAGAG 1524
QY 1659 CGGACACCTTCGCAAACTGGCCGAGAGAGTGG 1692
DB 1525 CGGCCACCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1558

RESULT 14
PCT-US93-00445-3
Sequence 3, Application PC/TUS9300445
GENERAL INFORMATION:
APPLICANT: Bell, Leonard
APPLICANT: Madril, Joseph A.
APPLICANT: Warren, Stephen L.
APPLICANT: Luthinger, Daniel J.
TITLE OF INVENTION: Genetically Engineered
TITLE OF INVENTION: Endothelial Cells
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 760 Kb storage
COMPUTER: DELL 486/50
OPERATING SYSTEM: DOS 5.0
SOFTWARE: Displaywrite 3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00445
FILING DATE: 19930105
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/820,011
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-101PCT
TELEPHONE: (203) 255 1400
TELEFAX: (203) 254 1101
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1611
TYPE: NUCLEIC ACID
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapien
POSITION IN GENOME:
CHROMOSOME/SEGMENT: Chromosome 20
PUBLICATION INFORMATION:
AUTHORS: Anderson, Stephen K.
AUTHORS: Gibbs, Carol P.
AUTHORS: Tanaka, Akio
AUTHORS: Kung, Hsing-Jien
AUTHORS: Fujita, Donald J.
TITLE: Human Cellular src Gene:
TITLE: Nucleotide Sequence and Derived Amino
TITLE: Acid Sequence of the Region Coding for
TITLE: the Carboxy-Terminal Two-Thirds of

Query Match 10.0%; Score 199.2; DB 1; Length 1602;
Best Local Similarity 55.7%; Pred. No. 7.2e-32;
Matches 453; Conservative 0; Mismatches 343; Indels 18; Gaps 3;

```
OY 897 CACGGGACCAAGTCGGCCGAGAGAGCTGGCCAGGCGGGCTGGTACTGACCTGCAG 956
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 736 CCCAGCTCCAAAGCCCGAGACCAGGAGCTCGCCAGAGGACCGCTGGAAATCCCGGGGAG 795
OY 957 CATTGACATTGGGAGACAGATCGAGAGGAGATTGGAGCTGTCTCTGCAGGG---T 1013
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 796 TCCTGCGGCTGGAGGTAGAGTGGGCGAGGCTCTTTGGAGAGTCTGGATGGGGAGC 855
OY 1014 GAGTACTGGGGCAAAAGGTGGCCGTGAAGATATCAAG---TGTGATGTGACAGCCGAG 1070
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 856 TGGAGCGGACACACAGAGTGGCCATAAAGACTGAAAGCCCGGCAACATGTCGCCGGAG 915
OY 1071 GCCTTCCTGGAGAGACGGCCGCTATGACGAAGATGCAACAGAGACCTGGTGCCTC 1130
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 916 GCCTTCCTGGAGAGACCGCAAGTGAAGAGCTCGGCAATGAGAGCTGGTTGAGCTG 975
OY 1131 CTGGGCTGATCTCTGACCAAGGGGCTGTACATTGTATGAGACAGTGCAGCAAGGCAAC 1190
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 976 TACGAGAGTGTGTGGAAGAGCCCATCTACATCTGCTAGTACATGAGCAAGGGAGC 1035
OY 1191 CTGGTAACTTCTGCGGACCCGGGGTGAAGCCCTGTGAACACCGCTGAGCTCTGCAG 1250
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1036 CTCTGGAATTCTGAAAGGAGAGATGGGCAAGTACCTGCGGCTCCACAGCTGCTGAT 1095
OY 1251 TTTTCTGCAAGTGGCGGAGGGCATGAGTACTGAGAGAGCAAGAGCTTGTGACCCG 1310
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1096 ATGGCTGCTCAAGTATGCTCCGGCAAGGCTATGTGAGAGATGAGTACTGATGACCGA 1155
OY 1311 GACCTGGCCGCGCCGCAACATCTGCTGACAGAGACTGCTGGCCAGGTGACGACTTT 1370
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1156 GACCTGGCGGGGCGCCACATCTGCTGGGGGAGAACTGCTGTGCAAGTGGCTGACTTT 1215
OY 1371 GGCTGGCCAAAGCCGAGCGGAAGGGGCTAGACTCAAGCCGGC-----TGCCC 1418
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1216 GGGCTGGCACGCTCATCGAGGACAAAGATACAGCAGGCAAGTGCCAAAGTTCCC 1275
OY 1419 GTCAAGTGGAGGGCGCCGAGGCTCTCAAAACAGGGGAAGTTCACAGCAAGTGGATGTC 1478
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1276 ATCAAGTGGACAGCCCGAGGAGCCCTCTATAGCCGGTTCCATCAAGTGGATGTC 1335
OY 1479 TGGAGTTTGGGGTCTGCTGAGGAGGTCTTCATATGACAGGGCTCCGTACCTTAA 1538
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1336 TGGTCTTGGGATCTGCTGTAGCTGAGCTGACCAAGGGCGGGTGCATACCCAGGG 1395
OY 1539 ATGTCACTGAAGAAGGTGTCGAGAGCCGTGAGAAAGGGGTACCGCATGGAACCCCGAG 1598
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1396 ATGCTCAACAGGAGGTGCTGAGCAGGTGAGAGGGGCTACCGCATGCCCTGCCGCC 1455
OY 1599 GGGCTGCCAGGCGCCGAGAGTCTTCATGAGCAGTGTGGAGGAGAGCCCGCCGCGC 1658
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1456 GAGTCCCCGAGTCCCTGATGACTCATGTGCTCAATGTGCTGCGGAGGAGCCCTGAGAG 1515
OY 1659 CGGCCACCTTCCGCAAACTGGCCGAGAAAGCTGG 1692
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1516 CGGCCACCTTTGAGTACCTGCAAGGCTTCTGG 1549
```

Search completed: August 1, 2003, 19:44:23
Job time : 187 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:32:53 ; Search time 669 Seconds
(without alignments)
8070.069 Million cell updates/sec

Title: US-09-977-260-1
Perfect score: 2000
Sequence: 1 ctgcctcaagtgtgcagc.....attctaagactcttaaaaa 2000

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 252756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1998.4	99.9	2000	16	AAT00616
2	1963.4	98.2	1997	24	AA150631
3	1909.8	95.5	1987	19	AAV44497
4	1901.8	95.1	1942	16	AA084888
5	1377	68.8	1713	24	ABK88791
6	547.6	27.4	2187	21	AAZ28701
7	547.6	27.4	2187	22	AAH28358
8	547.6	21.87	24	ABK84602	

9	547.6	27.4	2187	24	ABL62918
10	547.6	27.4	2420	21	AAZ46489
11	547.6	27.4	2420	25	ACC50120
12	547.6	27.4	2466	22	AA193804
13	527.8	26.4	2442	24	ABK73326
14	455	22.8	16389	24	ABK88792
15	330.2	16.5	2591	23	ABH11799
16	213.6	10.7	3299	22	ABH18556
17	213.6	10.7	4466	24	ABN59752
18	212.4	10.6	1090	23	AAK87964
19	212.4	10.6	1611	14	AAQ46688
20	212.4	10.6	1611	24	ABK73325
21	212.4	10.6	1611	25	ABZ59382
22	212.4	10.6	1699	23	AAK87965
23	204.2	10.2	2282	24	ABK56203
24	202.6	10.1	2032	21	AAZ44491
25	199.2	10.0	1602	14	AAQ46687
26	197.6	9.9	1759	21	AAZ29700
27	197.6	9.9	1759	22	AAH28357
28	193.4	9.7	1626	25	ABZ59389
29	192.6	9.6	1926	24	ABK83940
30	192.6	9.6	2015	24	ABK83939
31	192.6	9.6	2015	24	ABL66673
32	191.4	9.6	3323	25	ABZ18519
33	191.4	9.6	3380	25	ABZ18520
34	191.4	9.6	5520	18	AAT61865
35	191.4	9.6	5527	24	ABN97215
36	186.2	9.3	1911	24	ABK63704
37	185	9.2	1574	21	AAZ86794
38	185	9.2	1574	22	AAD11845
39	185	9.2	3527	24	ABO76288
40	185	9.2	3527	23	AAK66986
41	185	9.2	7487	23	AAK92457
42	184.6	9.2	2354	24	ABK83948
43	184.6	9.2	2354	24	ABL68108
44	184.6	9.2	2433	24	AAK94859
45	184.4	9.2	1254	12	AAQ13983

ALIGNMENTS

RESULT 1	ID	AAAT00616	standard: CDNA; 2000 BP.
AC	AAAT00616		
AC	AAAT00616		
DT	26-MAR-1996	(first entry)	
DE	Megakaryocyte kinase MKK1 CDNA.		
XX			
XX			
KW	Megakaryocyte kinase-1; MKK1; cytoplasmic tyrosine kinase;		
KW	cellular signal transduction; leukæmia; myelosis; myelofibrosis;		
XX	gene therapy; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	258..1781	
FT		/*tag= a	
XX			
PN	W09529185-A1.		
XX			
PD	02-NOV-1995.		
XX			
PF	24-APR-1995;	95WO-US05008.	
XX			
PR	21-APR-1995;	95US-0426509.	
XX			
PR	22-APR-1994;	94US-0232545.	
XX			
PA	(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.		
	(SUGE-) SUGEN INC.		

Breast cancer rela
PKA substrate, Csk
Breast cancer asso
Human polynucleoti
DNA encoding human
Gene encoding huma
Drosophila melanog
Human CDNA sequenc
Novel human coding
DNA encoding novel
Human p60 c-src 9
DNA encoding human
Human src-c encodi
DNA encoding novel
CDNA encoding tumo
PKA substrate, Src
Chicken p60 c-src
Wild-type chicken
Nucleotide sequenc
Mouse src-c encodi
Human CDNA differe
Human CDNA differe
Lung cancer relate
Group III CDNA can
c-abl gene. Homo
Gene #3713 used to
Rat sequence diffe
Human protein kina
Human protein kina
Human protein kina
DNA encoding novel
DNA encoding novel
Human CDNA differe
Ovary cancer relat
Human DNA sequence
Lk gene fused wit

xx Gishizky M, Sures I, Ullrich A:
xx WPI: 1995-382959/49.
xx P-PSDB: AAR84181.
xx New poly:nucleotide(s) encoding megakaryocyte tyrosine kinase(s) -
xx PT used to develop prods. for the treatment and diagnosis of kinase
xx related signal transduction abnormalities.
xx
xx Claim 2: Fig 1A-C; 82pp; English.
xx
xx Overlapping cDNA clones were combined to obtain a sequence (AAT00616)
xx coding for human megakaryocyte kinase MKK1 (AAR84181). The clones
xx were isolated from a fetal brain library using degenerate primers
xx (AAT00614-15) based on conserved regions within the kinase domain of
xx receptor tyrosine kinases. MKK polynucleotides can be used in the
xx procdn. in host cells of recombinant MKK, and in the gene therapy of
xx diseases such as acute megakaryocytic leukemia, myelofibrosis and
xx acute megakaryocytic myelosis.

SO Sequence 2000 BP; 375 A; 631 C; 674 G; 320 T; 0 other;

Query Match 99.9%; Score 1998.4; DB 16; Length 2000;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1999; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCGCTCCAAAGTTGTGACAGCCGGGAGCCCTGGGGTGTGACAGCGGCTGGGAGGCC 60
DB 1 CTCGCTCCAAAGTTGTGACAGCCGGGAGCCCTGGGGTGTGACAGCGGCTGGGAGGCC 60
QY 61 TCCGAGGG 120
DB 61 TCCGAGGG 120
QY 121 AGGCTCGGTCCAGTGGACACCAGCTCCCTACCTCTGTGACAGCGGCTGGGAGGCC 180
DB 121 AGGCTCGGTCCAGTGGACACCAGCTCCCTACCTCTGTGACAGCGGCTGGGAGGCC 180
QY 181 GGGCATTTCCCAAGGTGACAGCTGTGACAGCTGTGACAGCTGTGACAGCTGTGACAG 240
DB 181 GGGCATTTCCCAAGGTGACAGCTGTGACAGCTGTGACAGCTGTGACAGCTGTGACAG 240
QY 241 TTTCTCTGTGG 300
DB 241 TTTCTCTGTGG 300
QY 301 GCTGTGATTTCTGTGAGAACTTCCCGGGGTGAGCCCGCTTCTCCGAGCCTGGGACC 360
DB 301 GCTGTGATTTCTGTGAGAACTTCCCGGGGTGAGCCCGCTTCTCCGAGCCTGGGACC 360
QY 361 CCCCTCCCGTCTAGGAGAGATCCCAAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 420
DB 361 CCCCTCCCGTCTAGGAGAGATCCCAAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 420
QY 421 CCAAATGCGAGCACACCCGCCCAAGCAGAGGAGGAGCTTCCGAAAGGGGAGAGTGG 480
DB 421 CCAAATGCGAGCACACCCGCCCAAGCAGAGGAGGAGCTTCCGAAAGGGGAGAGTGG 480
QY 481 TCACATCTCTGAGAGGCTGCGAGAACAGAGTGTACCGCTCAAGACACACAGAGT 540
DB 481 TCACATCTCTGAGAGGCTGCGAGAACAGAGTGTACCGCTCAAGACACACAGAGT 540
QY 541 GACATGAGGG 600
DB 541 GACATGAGGG 600
QY 601 CCNAAGTCAAGCTCATGCGTGTGTTCCAGGAGAAATCTCGGGCCAGAGAGCTGTACAG 660
DB 601 CCNAAGTCAAGCTCATGCGTGTGTTCCAGGAGAAATCTCGGGCCAGAGAGCTGTACAG 660
QY 661 AGCTGACGCTCCCGAGAGATGGGCTGTCTCTGTGTGGGAGTCCGGCCGACCCCGGG 720
DB 661 AGCTGACGCTCCCGAGAGATGGGCTGTCTCTGTGTGGGAGTCCGGCCGACCCCGGG 720

DB 661 AGCTGACGCTCCCGAGAGATGGGCTGTCTCTGTGTGGGAGTCCGGCCGACCCCGGG 720
QY 721 ACTAGTCTCTGTGCTGAGCTTTTGGCCGACAGTATCATACAGGCTGTGACAGCG 780
DB 721 ACTAGTCTCTGTGCTGAGCTTTTGGCCGACAGTATCATACAGGCTGTGACAGCG 780
QY 781 ACGGCCACCTCACAATGATGAGGCGGTGTTCTTGTGCAACCTCATGACATGTGTGAGC 840
DB 781 ACGGCCACCTCACAATGATGAGGCGGTGTTCTTGTGCAACCTCATGACATGTGTGAGC 840
QY 841 ATTACAGCAAGAGAGGCGGCTATCTGCAACCAAGCTGTGTGAGCAAGAGGAGAAACAG 900
DB 841 ATTACAGCAAGAGAGGCGGCTATCTGCAACCAAGCTGTGTGAGCAAGAGGAGAAACAG 900
QY 901 GGACCAAGTGGCGGAGAGAGAGCTGTGACAGGCGGCTGTGTTCTCAACCTGTGAGATT 960
DB 901 GGACCAAGTGGCGGAGAGAGAGCTGTGACAGGCGGCTGTGTTCTCAACCTGTGAGATT 960
QY 961 TGACATTTGGGAGCACAATGCGAGAGAGAGATTGGAGCTGTCTCAGGGTGTGATACC 1020
DB 961 TGACATTTGGGAGCACAATGCGAGAGAGAGATTGGAGCTGTCTCAGGGTGTGATACC 1020
QY 1021 TGGGCAAAAGGTGGCGGTGAGAAATATCAAGTGTGATGTGACAGCCGACGCTTCTCTGG 1080
DB 1021 TGGGCAAAAGGTGGCGGTGAGAAATATCAAGTGTGATGTGACAGCCGACGCTTCTCTGG 1080
QY 1081 ACGAGAGCGGCTGATGACGAAGATGCAACAGCAAGAACTGTGCTGTCTGCGGCGTGA 1140
DB 1081 ACGAGAGCGGCTGATGACGAAGATGCAACAGCAAGAACTGTGCTGTCTGCGGCGTGA 1140
QY 1141 TCTGTGCAACAGGGGCTTACATTGTCTATGAGACAGCTGTGACCAAGGGCACTGGTGA 1200
DB 1141 TCTGTGCAACAGGGGCTTACATTGTCTATGAGACAGCTGTGACCAAGGGCACTGGTGA 1200
QY 1201 TTTGCGGAGCCGGGGGGTGCAGACCTGTGTAACACCGTCAAGCTTCTGCTGTCTGTG 1260
DB 1201 TTTGCGGAGCCGGGGGGTGCAGACCTGTGTAACACCGTCAAGCTTCTGCTGTCTGTG 1260
QY 1261 ACGTGGCGAGGGGATGAGTACCTGTGAGAGCAAGAAAGCTTGTGACACCGGACCTGG 1320
DB 1261 ACGTGGCGAGGGGATGAGTACCTGTGAGAGCAAGAAAGCTTGTGACACCGGACCTGG 1320
QY 1321 CCCGCAACATCTGTGTCTGAGAGAGCTGTGTTGGCCAAAGTGTGACGACTTGTGCTG 1380
DB 1321 CCCGCAACATCTGTGTCTGAGAGAGCTGTGTTGGCCAAAGTGTGACGACTTGTGCTG 1380
QY 1381 AAGCCGAGCGGAAGGGGCTAGACTCAAGCCGGCTGCCCTCAAGTGTGAGAGCGCCGAG 1440
DB 1381 AAGCCGAGCGGAAGGGGCTAGACTCAAGCCGGCTGCCCTCAAGTGTGAGAGCGCCGAG 1440
QY 1441 CTCTCAAAACAGGGGAAGTTACCAAGCAAGTGTGATGTTTGGGGTCTGTCTCT 1500
DB 1441 CTCTCAAAACAGGGGAAGTTACCAAGCAAGTGTGATGTTTGGGGTCTGTCTCTCT 1500
QY 1501 GGGAGGCTTCTATATGAGAGGGGCTCCGTAACCTAAATGTACAGTAAAGAGTGTGCG 1560
DB 1501 GGGAGGCTTCTATATGAGAGGGGCTCCGTAACCTAAATGTACAGTAAAGAGTGTGCG 1560
QY 1561 AGGCGGTGAGAGAGGGGTACCGCATGGAACCCCGGAGGGCTGTCCAGGCGCCGTGACAG 1620
DB 1561 AGGCGGTGAGAGAGGGGTACCGCATGGAACCCCGGAGGGCTGTCCAGGCGCCGTGACAG 1620
QY 1621 TCTCTATGAGACAGCTCTGTGAGAGAGAGCCGCGCGGCGCACCTTCCGAAACTGG 1680
DB 1621 TCTCTATGAGACAGCTCTGTGAGAGAGAGCCGCGCGGCGCACCTTCCGAAACTGG 1680
QY 1681 CCGAGAGAGCTGGCGGGGAGCTAAGAGTGTGACAGTGTGACAGTGTGACAGTGTGACAG 1740
DB 1681 CCGAGAGAGCTGGCGGGGAGCTAAGAGTGTGACAGTGTGACAGTGTGACAGTGTGACAG 1740
QY 1741 ACGCCAGAGCTTCAACTGCGCCCGAAGCAGAGGCTGTGACAGTGTGACAGTGTGACAG 1800
DB 1741 ACGCCAGAGCTTCAACTGCGCCCGAAGCAGAGGCTGTGACAGTGTGACAGTGTGACAG 1800

QY	1801	TGGGCCCCAGAGGACCGAGAGAGTGTAAGTGGCGCGTGGGGGGCACTACCAAGGCCCAAGG	1860
Db	1801	TGGGCCCCAGAGGACCGAGAGAGTGTAAGTGGCGCGTGGGGGGCACTACCAAGGCCCAAGG	1860
QY	1801	TGGGCCCCAGAGGACCGAGAGAGTGTAAGTGGCGCGTGGGGGGCACTACCAAGGCCCAAGG	1860
Db	1801	TGGGCCCCAGAGGACCGAGAGAGTGTAAGTGGCGCGTGGGGGGCACTACCAAGGCCCAAGG	1860
QY	1861	AGGGTCCAGAGGGGGCAAGTCACTCCTCGTGATGCCACACAGCAGGGGGTGGGCCACAGTAGG	1920
Db	1861	AGGGTCCAGAGGGGGCAAGTCACTCCTCGTGATGCCACACAGCAGGGGGTGGGCCACAGTAGG	1920
QY	1861	AGGGTCCAGAGGGGGCAAGTCACTCCTCGTGATGCCACACAGCAGGGGGTGGGCCACAGTAGG	1920
Db	1861	AGGGTCCAGAGGGGGCAAGTCACTCCTCGTGATGCCACACAGCAGGGGGTGGGCCACAGTAGG	1920
QY	1921	GGCTCTGGGGGGCCCCGTGGACACCCCAAGACTGCGAAGATGATCGCCCGATAAAGACGG	1980
Db	1921	GGCTCTGGGGGGCCCCGTGGACACCCCAAGACTGCGAAGATGATCGCCCGATAAAGACGG	1980
QY	1921	GGCTCTGGGGGGCCCCGTGGACACCCCAAGACTGCGAAGATGATCGCCCGATAAAGACGG	1980
Db	1921	GGCTCTGGGGGGCCCCGTGGACACCCCAAGACTGCGAAGATGATCGCCCGATAAAGACGG	1980
QY	1981	ATTCTAAGGACTCTAAAAAA 2000	
Db	1981	ATTCTAAGGACTCTAAAAAA 2000	

```

RESULT 2
AAL50631
ID AAL50631 standard; DNA; 1997 BP.

```

DT 02-JAN-2003 (first entry)
VY

Human matk gene-related coding sequence.

Human; gene; ds; allergic disease; atopic dermatitis

OS Homo sapiens.

FH	Key	Location/Qualifiers
FH	Key	Location/Qualifiers

```
/*tag= a
```

XX

XX
11

[illegible]

XX
FE
01-MAR-2002; 2002MO-0F01910.
XX

XX
XX
XX

0970T900-30T00Z
'T00Z-WWW-TZ

PA (GENU-) GENOA RES INC.
PA (NIGE-) JAPAN GEN AGENCY NATION.

AA Sugita Y, Hetshi M, Kaday S, Gunji S, Tsujimoto G;
PI

XX
DR WPI: 2002-750572/81.

DR F-PSDB; AAOL3660.
XX

PT Examining allergic diseases by changes in expression level of mark gene
PT in peripheral monocytes as indication, also applicable in screening

XX PI compounds for treating allergic diseases e.g. atopic dermatitis -

PS - Disclosure; Page 48-52; 60pp; Japanese.
XX

CC The invention comprises a method of examining allergic diseases, the
CC method comprises comparing the expression level of the matk gene in a
CC biological sample of a patient and a healthy individual. The method ob
CC the invention is useful for examining/diagnosing allergic diseases -
CC particularly atopic dermatitis, the method of the invention is also
CC useful in screening candidate compounds for remedies. The present DNA
CC sequence encodes a human matk-related protein.

SQ Sequence 1997 BP; 372 A; 631 C; 676 G; 318 T; 0 other;

Query Match	98.2%	Score 1963.4	DB 24	Length 1997
Best Local Similarity	99.8%	Pred. No. 0		
Matches 1987	Conservative	0	Mismatches 1	Indels 3
				Gaps 2

QY	1	TTTCGCTCCAAATTGTTGACAGCCGGGACCGCCTCGGGGTTGTGAGCGGGCTGCGGGAGGCGC	60
Db	8	CTCGCTCCAAATTGTTGACCGGGGACCGCCTCGGGGTTGTGAGCGGGCTGCGGGAGGCGC	67
QY	61	TTCTGAGGGGCGGGGCGGG--GCGGCTCTGGGGGGCGCCCTTGAGCAGAAAAAGAAAGAA	118
Db	68	TTCTGAGGGGCGGGGCGGGGCGGGCTCTGGGGGGCGCCCTTGAGCAGAAAAAGAAAGAA	127
QY	119	CCAGGCTTGGTTCAGAGTGGACCCAGCTTCCTACCTCTGTGCCAGCGGCTGGGCTGTGG	178
Db	128	CCAGGCTTGGTTCAGAGTGGACCCAGCTTCCTACCTCTGTGCCAGCGACTGGGCTGTGG	187
QY	179	CAGGCGATTTCCAGAGCTTCCCGGACTGTGACCACTTGTGCTCAATGTGTGCTCTCACTGCGCTC	238
Db	188	CAGGCGATTTCCAGAGCTTCCCGGACTGTGACCACTTGTGCTCAATGTGTGCTCTCACTGCGCTC	247
QY	239	AGTTTCCCTCTGGGGGGGAGTGGGCGGGGCGAGGCGCTCTGTGGTTTCTTGCGGGGCAATTCA	298
Db	248	AGTTTCCCTCT--GGGGGGGAGTGGGCGGGGCGAGGCGCTCTGTGGTTTCTTGCGGGGCAATTCA	306
QY	299	CGGCTGTGATTTCTGCTGAGGAATTTCCCGGGGTGAGCCCGCGCTTCTCTGAGGCTTGGA	358
Db	307	CGGCTGTGATTTCTGCTGAGGAATTTCCCGGGGTGAGCCCGCGCTTCTCTGAGGCTTGGA	366
QY	359	CCCCCTTCCGCTTTCAGGCAGGATGTGCCACAGAGCGCTGGGGCCCCGGGCACTGATAT	418
Db	367	CCCCCTTCCGCTTTCAGGCAGGATGTGCCACAGAGCGCTGGGGCCCCGGGCACTGATAT	426
QY	419	CACCAAAATGCGAGCACACCCGCCCAACAGCAGGGGAGCTGTGCTTCCGCAAGGGCGAGCT	478
Db	427	CACCAAAATGCGAGCACACCCGCCCAACAGCAGGGGAGCTGTGCTTCCGCAAGGGCGAGCT	486
QY	479	GGTCAACATCTCTGAGGCGCTCGAGAACAGAGCTGTGATCCGCGTCAAGCACACACACAG	538
Db	487	GGTCAACATCTCTGAGGCGCTCGAGAACAGAGCTGTGATCCGCGTCAAGCACACACACAG	546
QY	539	TGAGCAGAGAGGGGCTGTGGCAGCTGGGGGCGCTCGGGAGCGGGAGGCGCTTCTCCGAGA	598
Db	547	TGAGCAGAGAGGGGCTGTGGCAGCTGGGGGCGCTCGGGAGCGGGAGGCGCTTCTCCGAGA	606
QY	599	CCCCAAGTCAAGCCTCATGCGCGGTTCACAGGGAAATCTCGGGCCAGAGGGCGTCA	658
Db	607	CCCCAAGTCAAGCCTCATGCGCGGTTCACAGGGAAATCTCGGGCCAGAGGGCGTCA	666
QY	659	GCAGCTGCAGCCTCCCGAGAGATGGGCTTTCCTGGTGGGGAGTCCGCGCACACCCGG	718
Db	667	GCAGCTGCAGCCTCCCGAGAGATGGGCTTTCCTGGTGGGGAGTCCGCGCACACCCGG	726
QY	719	CGACTACGCTCTGTGAGCTTGTGGCGCGAGCTCATCCAGCTACCGGCTGTGCACCG	778
Db	727	CGACTACGCTCTGTGAGCTTGTGGCGCGAGCTCATCCAGCTACCGGCTGTGCACCG	786
QY	779	CGAGGGGCAACCTCAATCGATGAGGGCGGTTCCTTCGCAACCTCATGAGCAATGTGTGA	838
Db	787	CGAGGGGCAACCTCAATCGATGAGGGCGGTTCCTTCGCAACCTCATGAGCAATGTGTGA	846
QY	839	GCATTTCAGCAGAGCAAGAGGGCGCTATCTGCACCAAGCTGTGTGAGACCCAAAGCGGAACA	898
Db	847	GCATTTCAGCAGAGCAAGAGGGCGCTATCTGCACCAAGCTGTGTGAGACCCAAAGCGGAACA	906
QY	899	CGGCAACCAAGTGGCGCGAGAGAGAGCTGGCCAGGGCGGGCTGTACTGAACTGTGCAGA	958
Db	907	CGGCAACCAAGTGGCGCGAGAGAGAGCTGGCCAGGGCGGGCTGTACTGAACTGTGCAGA	966
QY	959	TTTGACATTTGGGAGACAGATCGGAGAGGGAGAGTTTGAGCTGTCCGAGGGGTGAGTA	1018
Db	967	TTTGACATTTGGGAGACAGATCGGAGAGGGAGAGTTTGAGCTGTCCGAGGGGTGAGTA	1028
QY	1019	CCTGGGGCAAAAGGTGGCGCTGTGAAGAAATCAAGTGTGATGTGACAGCCAGGCTTCTCT	1078
Db	1027	CCTGGGGCAAAAGGTGGCGCTGTGAAGAAATCAAGTGTGATGTGACAGCCAGGCTTCTCT	1086
QY	1079	GGAGGAGAGCGCCGTATGATGAGAGATGCAACAGAGAACCTTGGTGGCTTCTTGAGGCGT	1138

```

Db 1087 GGAGAGAGGCGCCCTCATGACGAGAGATGCAACACGAGACCTGGTGGCTCTCTGGGCGT 1146
Qy 1139 GATCTTCACACAGGGGGCTGTACATGTCATGAGACAGTGAAGGCAACCTGTGTAA 1198
Db 1147 GATCTTCACACAGGGGGCTGTACATGTCATGAGACAGTGAAGGCAACCTGTGTAA 1206
Qy 1199 CTTTCTGCGGACCCGGGGGTGAGCCCTGCGTGAACACCGCTCAGCTTCTCTCT 1258
Db 1207 CTTTCTGCGGACCCGGGGGTGAGCCCTGCGTGAACACCGCTCAGCTTCTCTCT 1266
Qy 1259 GCAGTGGCGGAGGAGCATGAGTACCTGAGAGAGCAAGAGTTGTGCACCGCGACCTGGC 1318
Db 1267 GCAGTGGCGGAGGAGCATGAGTACCTGAGAGAGCAAGAGTTGTGCACCGCGACCTGGC 1326
Qy 1319 CGCCCGCAACATCTGCTCTCAGAGACCTGTGGGCCAAGGTCAGCACTTTGGCTGGC 1378
Db 1327 CGCCCGCAACATCTGCTCTCAGAGACCTGTGGGCCAAGGTCAGCACTTTGGCTGGC 1386
Qy 1379 CAAAGCGGAGGAGGAGGCTAGACCTCAAGCCGGCTCCGTCAGTGGAGCGGCCGA 1438
Db 1387 CAAAGCGGAGGAGGAGGCTAGACCTCAAGCCGGCTCCGTCAGTGGAGCGGCCGA 1446
Qy 1439 GGCCTCTCAACACAGGGAATTCAACCAAGTCGATGTCTGAGAGTTTGGGGTGTCT 1498
Db 1447 GGCCTCTCAACACAGGGAATTCAACCAAGTCGATGTCTGAGAGTTTGGGGTGTCT 1506
Qy 1499 CTGGGAGGTTTCTCATATGAGACGGGCTCCGTACCTTAATATGCTACTGAAAGAGTGT 1558
Db 1507 CTGGGAGGTTTCTCATATGAGACGGGCTCCGTACCTTAATATGCTACTGAAAGAGTGT 1566
Qy 1559 GGAGGCGGTGAGAGAGGGGTACCGCAATGAAACCCCGGAGGGCTGTCCAGCCCGCTGCA 1618
Db 1567 GGAGGCGGTGAGAGAGGGGTACCGCAATGAAACCCCGGAGGGCTGTCCAGCCCGCTGCA 1626
Qy 1619 CGTCTCATGAGAGCTGTGTGGAGGAGAGCCCGCGCGCCGACCTTCGCGAAT 1678
Db 1627 CGTCTCATGAGAGCTGTGTGGAGGAGAGCCCGCGCGCCGACCTTCGCGAAT 1686
Qy 1679 GGCGGAGAGAGCTGGCGGGAGAGTACGCAATGCAAGTGTCCCGACCTCGCTCAGGGCA 1738
Db 1687 GGCGGAGAGAGCTGGCGGGAGAGTACGCAATGCAAGTGTCCCGACCTCGCTCAGGGCA 1746
Qy 1739 GGAGGCGGAGGCTGTGTGGAGGAGAGCCCGCGCGCCGACCTTCGCGAAT 1798
Db 1747 GGAGGCGGAGGCTGTGTGGAGGAGAGCCCGCGCGCCGACCTTCGCGAAT 1806
Qy 1799 CTGGGCGGAGAGAGAGTGAAGTGTGGGGTGTGGGCACTGACAGGCCAA 1858
Db 1807 CTGGGCGGAGAGAGAGTGAAGTGTGGGGTGTGGGCACTGACAGGCCAA 1866
Qy 1859 GGAAGGCTCCAGGGGGCAAGTATCTCTCTGGTGGCCACAGAGGGGCTGGCCACGTAG 1918
Db 1867 GGAAGGCTCCAGGGGGCAAGTATCTCTCTGGTGGCCACAGAGGGGCTGGCCACGTAG 1926
Qy 1919 GGGGCTCTGGGGCGCCGTGGAGACACCCAGCACTGCGAAGGATGCGCCGATAAAG 1978
Db 1927 GGGGCTCTGGGGCGCCGTGGAGACACCCAGCACTGCGAAGGATGCGCCGATAAAG 1986
Qy 1979 GGATTTCTAAGG 1989
Db 1987 GGATTTCTAAGG 1997

```

RESULT 3
AAV44497
ID AAV44497 standard; cDNA; 1987 BP.
XX
AC AAV44497;
XX
DT 16-OCT-1998 (first entry)
XX
DE Human matk cDNA.

```

XX XX Breast; cancer; matk; CSK homologous kinase; CHK; detection; diagnosis;
KW cytoplasmic protein; tyrosine kinase; ErbB-2; negative regulator;
KW mitogenic signalling; ss.
OS Homo sapiens.
XX Key Location/Qualifiers
FH 263..1846
FT CDS /*tag= a
FT /product= CHK
FT /note= "CSK homologous kinase"
PE W09830704-A1.
PD 16-JUL-1998.
PD 07-JAN-1998; 98MO-US00420.
PR 16-JUN-1997; 97US-0876882.
PR 08-JAN-1997; 97US-0035228.
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PI Avraham H, Groopman JE;
PI WPI: 1998-399149/34.
PI P-PSDB: AAM64454.
PT Detecting breast cancer by detecting Csk homologous kinase
PT expression - especially in humans and use of Csk homologous kinase
PT in treatment or prophylaxis of breast cancer and for producing
PT medications
PS Disclosure; Fig 2; 54pp; English.
XX This sequence encodes a CSK homologous kinase (CHK) which is used in a
XX method of detecting cancer in breast tissue. The method allows diagnosis
XX of breast cancer in mammals, especially humans. It is based on the
XX discovery that a cytoplasmic protein tyrosine kinase, CHK, is expressed
XX in human breast tissue, but not in adjacent tissue. This protein can be
XX used to raise antibodies which can be included in compositions and
XX diagnostic kits for diagnosis of breast cancer. The presence of CHK in
XX breast tissue can also be determined using other standard methods (e.g.
XX Northern blotting) or by detecting nucleic acid sequences encoding all/a
XX portion of the protein (e.g. using hybridisation probes). Over-expression
XX of the receptor tyrosine kinase ErbB-2 has previously been associated
XX with the development of breast cancer, and CHK specifically interacts
XX with activated ErbB-2, and may function as a negative regulator of
XX ErbB-2 mediated mitogenic signalling. The compositions may also be used
XX to design drugs (e.g. which incorporate CHK analogues with greater
XX biological activity than CHK) and to identify CHK antagonists and
XX agonists for therapeutic use.
XX
XX Sequence 1987 BP; 369 A; 628 C; 672 G; 318 T; 0 other;
Qy 1 CTCGCTCCAAAGTGTGTCAGACCGGAGCGCTCGGGGTGTGACACCGGCTCGGAGAGCC 60
Db 8 CTCGCTCCAAAGTGTGTCAGACCGGAGCGCTCGGGGTGTGACACCGGCTCGGAGAGCC 67
Qy 61 TCCTGGGGGCGGGGCGGCGCTCGGGGGGCGCCCTGAGAGCAAGAAACAGGAAGAAC 120
Db 68 TCCTGGGGGCGGGGCGGCGCTCGGGGGGCGCCCTGAGAGCAAGAAACAGGAAGAAC 127
Qy 121 AGGCTGGTCCAGTGGACACCAAGCTTCTTACCTCTGTGTGCAAGCCGCTGTGGCA 180
Db 128 AGGCTGGTCCAGTGGACACCAAGCTTCTTACCTCTGTGTGCAAGCCGCTGTGGCA 187
Qy 181 GGCAATTCCAGCGTCCCGAGCTGTGACCACTTGTCTGAGTGTGCTTCACTCGCTCAG 240

```

Db 188 GGCATTCCACGCGCCGACCTGTGACACTTGTGCTCAGTGTGCTCTACCTGCTCAG 247
 Oy 241 TTTCCTCTGGGGGCGATGGGGGGGCGAGGCTCTGTGTTCCGTGGGGGATTTACG 300
 Db 248 TTTCC - -TCGCGGGGCGATGGGGGGCGAGGCTCTGTGTTCCGTGGGGGATTTACG 305
 Oy 301 GCTGTGATTTCTGTGAGGAACTTCCCGGGGTGAGCCCGGCTTCTCCGAGCCTGGAC 360
 Db 306 GCTGTGATTTCTGTGAGGAACTTCCCGGGGTGAGCCCGGCTTCTCCGAGCCTGGAC 365
 Oy 361 CCCCTCCGCTCAAGCAGATGCCAGAGGCGCTGGGGCCCGGGGCAACCAAGTATCA 420
 Db 366 CCCCTCCGCTCAAGCAGATGCCAGAGGCGCTGGGGCCCGGGGCAACCAAGTATCA 425
 Oy 421 CCAATGGAGACACCGCCCGCAGCGAGGAGCTGGCTTCCGCAAGGGCCAGCTGG 480
 Db 426 CCAATGGAGACACCGCCCGCAGCGAGGAGCTGGCTTCCGCAAGGGCCAGCTGG 485
 Oy 481 TCACATCTGGAGGCTGTGCGAACAAGAGCTGTACCGGTCAAGCACCACAGTGG 540
 Db 486 TCACATCTGGAGGCTGTGCGAACAAGAGCTGTACCGGTCAAGCACCACAGTGG 545
 Oy 541 GACAGAGGCGCTGTGCGACAGTGGGGCGCTGTGCGAGCGGGAGCGGCTTCCGAGAC 600
 Db 546 GACAGAGGCGCTGTGCGACAGTGGGGCGCTGTGCGAGCGGGAGCGGCTTCCGAGAC 605
 Oy 601 CCAAGCTAGCTCTAGCGGTGTCCAGAGGAAATCTCCGGCCAGAGAGGTGTCCAGC 660
 Db 606 CCAAGCTAGCTCTAGCGGTGTCCAGAGGAAATCTCCGGCCAGAGAGGTGTCCAGC 665
 Oy 661 AGCTGACAGCTCCCGAGAGTGGGCTGTCTGTGCGAGAGTCCGCGCCACCCGCGG 720
 Db 666 AGCTGACAGCTCCCGAGAGTGGGCTGTCTGTGCGAGAGTCCGCGCGCCACCCGCGG 725
 Oy 721 ACTAGCTCTGTGCTGAGCTTTTGGCGCGAGCTATCCACTACCGCTGTGACCGCG 780
 Db 726 ACTAGCTCTGTGCTGAGCTTTTGGCGCGAGCTATCCACTACCGCTGTGACCGCG 785
 Oy 781 ACGGCACTCTCAATCATGTAGGCGGCTTGTGCACTCATGTAGCAATGTGTGAGC 840
 Db 786 ACGGCACTCTCAATCATGTAGGCGGCTTGTGCACTCATGTAGCAATGTGTGAGC 845
 Oy 841 ATTACAGCAAGACAAAGGCGCTATCTGACCAAGCTGTGAGACCAAGGGAACAGC 900
 Db 846 ATTACAGCAAGACAAAGGCGCTATCTGACCAAGCTGTGAGACCAAGGGAACAGC 905
 Oy 901 GGACCAAGTCCGCGAGAGAGAGTGGCGAGGCGGCTGTGTAAGTGAAGTGAAGT 960
 Db 906 GGACCAAGTCCGCGAGAGAGAGTGGCGAGGCGGCTGTGTAAGTGAAGTGAAGT 965
 Oy 961 TGACATTGGAGACACAGATGGAGAGAGAGTGGAGCTGTGCAAGGTTGAGTACC 1020
 Db 966 TGACATTGGAGACACAGATGGAGAGAGAGTGGAGCTGTGCAAGGTTGAGTACC 1025
 Oy 1021 TGGGGCAAAAGTGGCCGTGAAGATATCAAGTGTGATGTACAGCCAGGCTTCTCTG 1080
 Db 1026 TGGGGCAAAAGTGGCCGTGAAGATATCAAGTGTGATGTACAGCCAGGCTTCTCTG 1085
 Oy 1081 ACGAGAGCGGCGCTCATGACGAAGATGCAACAGAGACCTGTGCTTCTGGGCGTGA 1140
 Db 1086 ACGAGAGCGGCGCTCATGACGAAGATGCAACAGAGACCTGTGCTTCTGGGCGTGA 1145
 Oy 1141 TCCGACACGAGGGGCTGTACATTTGATGAGACAGCTGAGCAAGGGCAACCTGGGACT 1200
 Db 1146 TCCGACACGAGGGGCTGTACATTTGATGAGACAGCTGAGCAAGGGCAACCTGGGACT 1205
 Oy 1201 TTCTGCGAGCCCGGGGCTGAGCCCTGTGTAACACCGCTCAGCTCTGCAAGTTTCTGCG 1260
 Db 1206 TTCTGCGAGCCCGGGGCTGAGCCCTGTGTAACACCGCTCAGCTCTGCAAGTTTCTGCG 1265
 Oy 1261 ACGTGGCGGAGGAGTGAAGTACTGTGAGAGCAGAAAGCTTGTGACCGGCACTGGCGG 1320

Db 1266 ACGTGGCGGAGGGAGTGAAGTACTGAGAGCAAGAAAGCTTGTGACCCGCACTGGCG 1325
 Oy 1321 CCCGCAATCTGTGCTCAGAGGACCTGTGTGCCAAGTCAAGCACTTTGGCCCTGGCA 1380
 Db 1326 CCCGCAATCTGTGCTCAGAGGACCTGTGTGCCAAGTCAAGCACTTTGGCCCTGGCA 1385
 Oy 1381 AAGCGAGCGGAGAGGGCTAGACTCAACCGGCTCCCGTCAAGTGAAGCGGCGCCGAG 1440
 Db 1386 AAGCGAGCGGAGAGGGCTAGACTCAACCGGCTCCCGTCAAGTGAAGCGGCGCCGAG 1445
 Oy 1441 CTCTCAACACGCGGAGTTCAACAGCAAGTGTGTGAGTTTGGGGTGTGCTCT 1500
 Db 1446 CTCTCAACACGCG - -GTTCAACAGCAAGTGTGTGAGTTTGGGGTGTGCTCT 1502
 Oy 1501 GGGAGCTTCTCATATGAGAGGCTCCGTACCTTAAATGTCACTGAAGAGTGTGG 1560
 Db 1503 GGGAGCTTCTCATATGAGAGGCTCCGTACCTTAAATGTCACTGAAGAGTGTGG 1562
 Oy 1561 AGGCGTGGAGAGGGGTACCGCATGGAACCCCGAGAGGCTGTCCAGGCGCCGTGACG 1620
 Db 1563 AGGCGTGGAGAGGGGTACCGCATGGAACCCCGAGAGGCTGTCCAGGCGCCGTGACG 1622
 Oy 1621 TCCTCATAGCAGCTGTGTGGAGGAGAGCCCGCGCGCACCTTCCGCAAACTGG 1680
 Db 1623 TCCTCATAGCAGCTGTGTGGAGGAGAG - CGCGCGCGCGCACCTTCCGCAAACTGG 1681
 Oy 1681 CCGAAGACTGGCCCGGAGAGTACCGCATGGAAGTGTGCCCAAGCTTCTTGAAGGCAAG 1740
 Db 1682 CCGAAGACTGGCCCGGAGAGTACCGCATGGAAGTGTGCCCAAGCTTCTTGAAGGCAAG 1741
 Oy 1741 ACGCGAGCGGCTCACTCGCCCGCAGAGCAGAGACCTGTGACCCAGCGGCGGCGCT 1800
 Db 1742 ACGCGAGCG - TCCACTCGCCCGCAGAGCAGAGACCTGTGACCCAGCGGCGGCT 1798
 Oy 1801 TGGCCCGAGAGAGCCGAGAGAGTGAAGTGTGGCGGCTGTGACAGAGCCCAAG 1860
 Db 1799 TGGCCCGAGAGAGCCGAGAGAGTGAAGTGTGGCGGCTGTGACAGAGCCCAAG 1858
 Oy 1861 AGGCTCAGAGGCGGCAAGTCACTCTGTGTGCCACAGAGGCGCTGGCCCACTAGG 1920
 Db 1859 AGGCTCAGAGGCGGCAAGTCACTCTGTGTGCCACAGAGGCGCTGGCCCACTAGG 1918
 Oy 1921 GGCCTTGGGCGGCGCGTGGAGACCCAGACCTGTGCAAGAGTATGCCGATTAAGACG 1980
 Db 1919 GGCCTTGGGCGGCGCGTGGAGACCCAGACCTGTGCAAGAGTATGCCGATTAAGACG 1978
 Oy 1981 ATTCTAAG 1989
 Db 1979 ATTCTAAG 1987
 RESULT 4
 AA084888
 ID AA084888 standard; cDNA to mRNA; 1942 BP.
 XX
 AC AA084888;
 XX
 DT 25-MAR-2003 (updated)
 DT 26-OCT-1995 (first entry)
 DE DNA encoding cytoplasmic tyrosine kinase.
 KW cytoplasmic; tyrosine kinase; blood; cell differentiation;
 KW screening; anticancer agent; ds.
 OS Homo sapiens.
 XX
 XX
 FH Key location/Qualifiers
 FT CDS 208..1731
 FT /tag= a
 FT /product= cytoplasmic-tyrosine_kinase
 FT /note= "see AAR71133"
 FT misc_feature 349..540

FT FT /tag- b
FT /note- "encodes SH3 domain (see AAR71129)"
FT misc_feature 571..795
FT /tag- c
FT /note- "encodes SH2 domain (see AAR71130)"
FT misc_feature 904..1641
FT /tag- d
FT /note- "encodes tyrosine kinase domain (see AAR71131)"
FT misc_feature 331..1728
FT /tag- e
FT /note- "encodes N-terminally truncated form of the enzyme (see AAR71132)"
XX WO9506113-A1.
XX 02-MAR-1995.
XX 25-AUG-1994.
XX 94WO-JP01411.
XX 25-AUG-1993.
XX 93JP-0210403.
XX 29-MAR-1994.
XX 94JP-0058553.
XX (ASAH) ASAHU KASEI KOGYO KK.
XX Sakano S.
PI WPI: 1995-106842/14.
DR P-PSDB: AAR71129, AAR71130, AAR71131, AAR71132, AAR71133.
XX Cytoplasmic tyrosine kinase and antibody recognising it - for
PT screening chemical substances for tyrosine kinase inhibitory or
XX activating activity for use as cancer therapy
XX
XX Claim 7: Page 49-50; 58pp; English.
XX
XX This DNA encodes a cytoplasmic tyrosine kinase which has enhanced
CC expression in connection with blood cell differentiation. It was
CC isolated from the human UT-7 blood cell line. The DNA sequences and
CC antibodies raised against the enzyme, are useful for screening agents
CC for inhibiting or activating activity on the tyrosine kinase, for
CC use as anticancer agents.
CC (updated on 25-MAR-2003 to correct PN field.)
CC
SQ Sequence 1942, BP: 365 A; 615 C; 651 G; 311 T; 0 other;
Query Match 95.1%; Score 1901.8; DB 16; Length 1942;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1938; Conservative 0; Mismatches 2; Indels 5; Gaps 3;
QY 52 CGGAGGCGCTCTGCGGGGGGGGGGGG--GCGGCTCGGGGGGGGGGGGGGAGAGAGAAA 109
DB 1 CGGAGGCGCTCTGCGGG 60
QY 110 CAGGAGAACCGAGGCGGCTGCGAGTGGGACCCAGCTCCCTACTCTCTGTGCGACCGGCT 169
DB 61 CAGGAGAACCGAGGCGGCTGCGAGTGGGACCCAGCTCCCTACTCTCTGTGCGACCGGCT 120
QY 170 GGGCTGTGGAGGCGATTCACAGCGTCCCGAGTGTGACCACTTGTGCTAGTGGCTTC 229
DB 121 GGGCTGTGGAGGCGATTCACAGCGTCCCGAGTGTGACCACTTGTGCTAGTGGCTTC 180
QY 230 ACCGTGCTGAGTTCCTCTGCGGGGGGCGATGGCGGGGGCGAGGCTCTGTGTTTCCGCGG 289
DB 181 ACCGTGCTGAGTTCCTCTCT- GGGGGGCGATGGCGGGGGCGAGGCTCTGTGTTTCCGCGG 239
QY 290 GGCATTTCAGCGGTGTGATTCCTGAGTGAAGAACTTCCCGGGGTAGCGCCCGCTTCTCGG 349
DB 240 GGCATTTCAGCGGTGTGATTCCTGAGTGAAGAACTTCCCGGGGTAGCGCCCGCTTCTCGG 299
QY 350 AGCGTGGCAAGCGGCTCCGCTCAGCGAGATGCAAGAGAGGCGGTGGGCGCGCGGCGAC 409
DB 300 AGCGTGGCAAGCGGCTCCGCTCAGCGAGATGCAAGAGAGGCGGTGGGCGCGCGGCGAC 359

QY 410 CCAAGTATACCAAAATGCGAGACACCCCGCCCAAGCCAGGGAGCTGGCTTCCGCA 469
DB 360 CCAAGTATACCAAAATGCGAGACACCCCGCCCAAGCCAGGGAGCTGGCTTCCGCA 419
QY 470 GGGGAGCGTGTACCACTCTGAGGCTCTCGAGAGACAGAGCTGTACCGCTCAAGCA 529
DB 420 GGGGAGCGTGTACCACTCTGAGGCTCTCGAGAGACAGAGCTGTACCGCTCAAGCA 479
QY 530 CCAACACAGTGAAGAGAGGGGGGCTGGAGCTGGGGGGGGGGGGGGGGGGGGGGGGGG 589
DB 480 CCAACACAGTGAAGAGAGGGGGGCTGGAGCTGGGGGGGGGGGGGGGGGGGGGGGGGG 539
QY 590 CTCGCGAG 649
DB 540 CTCGCGAG 599
QY 650 GGGCTGTACAG 709
DB 600 GGGCTGTACAG 659
QY 710 CCAAGCGGAG 769
DB 660 CCAAGCGGAG 719
QY 770 GCTGCAAGCGAG 829
DB 720 GCTGCAAGCGGAG 779
QY 830 CATGTGTGAG 889
DB 780 CATGTGTGAG 839
QY 890 GCGGAAACAGGAG 949
DB 840 GCGGAAACAGGAG 899
QY 950 CCGTGAAGATTTGACATTTGGAGACACAGATCGAGAGAGAGAGAGAGAGAGAGAGAG 1009
DB 900 CCGTGAAGATTTGACATTTGGAGACACAGATCGAGAGAGAGAGAGAGAGAGAGAGAG 959
QY 1010 GGGTGAAGATCTGGGGGCAAAAGTGGCGGTGAAGATATATGATGATGACACCCA 1069
DB 960 GGGTGAAGATCTGGGGGCAAAAGTGGCGGTGAAGATATATGATGATGACACCCA 1019
QY 1070 GGGCTTCTGAG 1129
DB 1020 GGGCTTCTGAG 1079
QY 1130 CCGTGGCGGTGATCTGCACAGAGGGGCTGTACATTTGTATGAGACAGTGAAGGGCAA 1189
DB 1080 CCGTGGCGGTGATCTGCACAGAGGGGCTGTACATTTGTATGAGACAGTGAAGGGCAA 1139
QY 1190 CCGTGGTGAATCTTCTCGAGACCCGGGGGTGAGAGCCCTGCTGTGAACACCGCTCAGT 1249
DB 1140 CCGTGGTGAATCTTCTCGAGACCCGGGGGTGAGAGCCCTGCTGTGAACACCGCTCAGT 1199
QY 1250 GTTTTCTGACAGTGGCGAGAGGCGATGAGTACCTGAGAGAGAGAGAGAGAGAGAGAGAG 1309
DB 1200 GTTTTCTGACAGTGGCGAGAGGCGATGAGTACCTGAGAGAGAGAGAGAGAGAGAGAGAG 1259
QY 1310 CGAGCTGGCGCGCCGAGACATCTGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1369
DB 1260 CGAGCTGGCGCGCCGAGACATCTGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1319
QY 1370 TGGCTTGGCCAAAGCGAG 1429
DB 1320 TGGCTTGGCCAAAGCGAG 1379
QY 1430 GGGCGCCGAGAGCTCTCAAAACAGGGAAGTTTCAACAGAGAGAGAGAGAGAGAGAGAG 1489
DB 1380 GGGCGCCGAGAGCTCTCAAAACAGGGAAGTTTCAACAGAGAGAGAGAGAGAGAGAGAG 1439

QY	1490	GGTCCTGCTGGGAGGCTCTCATATATGACACGGGCTCCGTAACCTTAATATGACGAA	154.9
Db	1440	GGTCTGCTCTGGGAGGCTCTTCTATATGACGGGCTCCGTAACCTTAATATGACGAA	149.99
QY	1550	AGAGGTGTCGGAGGCCCTGGAGAAAGGGGTACCCGATGGAAATCCCCCGAGGGCTGTCCAG	160.99
Db	1500	AGAGGTGTCGGAGGCCCTGGAGAAAGGGGTACCCGATGGAAATCCCCCGAGGGCTGTCCAG	155.99
QY	1610	CCCGGTGACAGTCCCTCATATGACACTGCTGGGAGGACAGACGCCGCCGCCGACACCTT	166.69
Db	1560	CCCGGTGACACTCTCTATGACACTGCTGGGAGGACAGACGCCGCCGCCGACACCTT	161.99
QY	1670	CCGCAAACTGGCCACAGAACTGTGGCCCGGAGACTACGAGTGCAGGTGCCCGACCTCCGT	172.99
Db	1620	CCGCAAACTGGCCAGAAAGCTGGCCCGGAGACTACGAGTGCAGGTGCCCGACCTCCGT	167.9
QY	1730	CTCAGGGCAGAGACGCCCGACGGCTCTCACTGGCCCCGAGAACCGAGAGCCCTTACCCACCC	178.99
Db	1680	CTCAGGGCAGAGACGCCCGACGGCTCTCACTGGCCCCGAGAACCGAGAGCCCTTACCCACCC	173.99
QY	1790	GGTGGGGCCCTTGGGCCCCACAGAGACCGAGAGTGTGAGAGTGCAGGCGTGGGGGACACTGAC	184.9
Db	1740	GGT--GGGCCCTTGGGCCCCACAGAGACCGAGAGTGTGAGAGTGTGGGGGACACTGAC	179.97
QY	1850	CAGGCCCAAGAGGAGGTCCAGGCGGGGCAAGTATCTCTGCTGCTGCCACAGCAGGAGCTGG	190.99
Db	1798	CAGGCCCAAGAGGAGGTCCAGGCGGGGCAAGTATCTCTGCTGCTGCCACAGCAGGAGCTGG	185.79
QY	1910	CCCAAGTAAAGGGGCTGTGGGCGGGCCGTGAGACCCAGAGCTCGAGAAAGTGTGATGGCCC	196.99
Db	1858	CCCAAGTAAAGGGGCTGTGGGCGGGCCGTGAGACCCAGAGCTCGAGAAAGTGTGATGGCCC	191.7
QY	1970	GATTAAGACGAGATTCTAAGGACTCT	199.4
Db	1918	GATTAAGACGAGATTCTAAGGACTCT	194.2

RESULT 5	
ABK88791	
ID	ABK88791 standard; cDNA; 1713 BP.
XX	
AC	
ABK88791;	
XX	
DT	21-OCT-2002 (first entry)
XX	
DE	cDNA encoding human kinase related to tyrosine kinase family
XX	
KW	Human; kinase; tyrosine kinase; kinase mediated disorder;
KW	haematopoietic disorder; cell signal transduction disorder;
KW	cancer; haemostatic; cytostatic; chromosome 19; gene; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	223..1470
CDS	/tag= a
FT	/product= "Kinase"
XX	
PN	W0200252018-A2.
XX	
PD	04-JUL-2002.
XX	
PF	19-DEC-2001; 2001MO-US48546.
XX	
PR	21-DEC-2000; 2000US-0741154.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	Beasley EM, Shao W, Ketchum K, Di Francesco V;
XX	
WI	PI; 2002-583568/62.
DR	P-PSDB; AAU9575.
DR	
XX	

PT	New isolated human kinase proteins and genes, useful in developing
PT	drugs, as well as for diagnosing, preventing or treating disorders
PT	associated with defective cell signal transduction, e.g. cancer or
PT	haematopoietic disorders
PS	
XX	Claim 4; Fig 1A; 78pp; English.
CC	The present invention relates to the isolation of a novel human kinase
CC	related to the tyrosine kinase family, and polynucleotide sequences
CC	encoding it. The gene encoding the kinase of the invention maps to
CC	chromosome 19. The kinase may be used for identifying a modulator of
CC	the kinase, an agent that binds to the kinase, or for identifying
CC	other members of the family. The kinase may also be used to raise
CC	antibodies which may be used in immunoassays or drug screening assays.
CC	The sequences of the invention may be used for treating a disease or
CC	condition mediated by a human kinase such as haematopoietic disorders
CC	and other disorders associated with defective cell signal transduction
CC	(e.g. cancer). The polynucleotide sequences encoding the kinase are
CC	useful for isolating and purifying the kinase, and as probes or primers.
CC	The present sequence encodes the human kinase of the invention.
XX	
SO	Sequence 1713 BP; 374 A; 499 C; 575 G; 265 T; 0 other:
Query Match	68.8%; Score 1377; DB 24; Length 1713;
Best Local Similarity	99.3%; Pred. No. 2.1e-266;
Matches 1383; Conservative	0; Mismatches 10; Indels 0; Gaps 0;
OY	608 CAGCCTCATGCCGNGTTCACAGGGAACATCTCGGGCCAGAGAGGTGTCCACGACTGCA 667
DB	297 CAGCAGCTTTGGTGGTTCACAGGGAACATCTCGGGCCAGAGAGGTGTCCACGACTGCA 356
OY	668 GCCTCCGAGGATGGGCTTCTCTGGTCCGGGAGTCCGCGGCCACCCCGGAGTACGT 727
DB	357 GCCTCCGAGGATGGGCTTCTCTGGTCCGGGAGTCCGCGGCCACCCCGGAGTACGT 416
OY	728 CCTGTGGGTGAGCTTTGGCCCGAGCGTATCCACTACCCGCTGTGCAACCGGACGCCA 787
DB	417 CCTGTGGGTGAGCTTTGGCCCGAGCGTATCCACTACCCGCGGTGTGCAACCGGACGCCA 476
OY	788 CCTCACAATGCATGAGGCGCGTCTTCTGTGCAACCTCATGGACATGAGGAGCATTTACAG 847
DB	477 CCTCACAATGCATGAGGCGCGTCTTCTGTGCAACCTCATGGACATGAGGAGCATTTACAG 536
OY	848 CAAGGACAAAGGCGCTATCTCACCACAGCTGGTAGACCAAAAGCGGAACACGGGACCAA 907
DB	537 CAAGGACAAAGGCGCTATCTCACCACAGCTGGTAGACCAAAAGCGGAACACGGGACCAA 596
OY	908 GTGGCGCAGAGAGAGAGCTGGCCAGGGCGGCTGGTTACTGAACCTGCAAGCTTTTACATT 967
DB	597 GTGGCGCAGAGAGAGAGCTGGCCAGGGCGGCTGGTTACTGAACCTGCAAGCTTTTACATT 656
OY	968 GGGAGACAGATCGGAGAGGAGAGATTGGAGCTGTCTCTGCAAGGTTGAGTACCTGGGGCA 1027
DB	657 GGGAGACAGATCGGAGAGGAGAGATTGGAGCTGTCTCTGCAAGGTTGAGTACCTGGGGCA 716
OY	1028 AAAGGTGGCCGTGAAGAATATCAAGTGTGATGTGACAGCCGACAGGCGCTTCTGTGAGAGAC 1087
DB	717 AAAGGTGGCCGTGAAGAATATCAAGTGTGATGTGACAGCCGACAGGCGCTTCTGTGAGAGAC 776
OY	1088 GGGCGTCATGACGAAGATGCAACACGAAACCTGGTGGCTCTCCGCGGCGTGATCTTCA 1144
DB	777 GGGCGTCATGACGAAGATGCAACACGAAACCTGGTGGCTCTCCGCGGCGTGATCTTCA 836
OY	1148 CCAGGGGCTGTACATTGTTCATGAGACAGTGAAGGCAACCGTGTGAACCTTCTGCG 1207
DB	837 CCAGGGGCTGTACATTGTTCATGAGACAGTGAAGGCAACCGTGTGAACCTTCTGCG 896
OY	1208 GACCCGGGGGTGAAGCCTCTGTGAACACCGCTCAGCTCTGTGCAAGTTTCTGTGACGTGGC 1267
DB	897 GACCCGGGGGTGAAGCCTCTGTGAACACCGCTCAGCTCTGTGCAAGTTTCTGTGACGTGGC 956
OY	1268 CGAGGGCATGAGTACCTCGAGAGCAAGAACTTGTGACAGCGGCAACCTGGCGGCCGCA 1327

```

Db 957 CGAGGGCATGAGTACCTGGAGAGCAAGAAGCTTGTGACCGCGACCTGGCCGCCGCA 1016
Qy 1328 CATCTGTGCTCACAAGAGCTGGTGGCCAAAGTCAAGCACTTTGGCTGGCCAAAGCGA 1387
Db 1017 CATCTGTGCTCACAAGAGCTGGTGGCCAAAGTCAAGCACTTTGGCTGGCCAAAGCGA 1076
Qy 1388 GCGAAGGGGCTAGACTCAAGCCGGCTGCAAGTGAAGGAGCGGCGCCGAGGCTCTCAA 1447
Db 1077 GCGAAGGGGCTAGACTCAAGCCGGCTGCAAGTGAAGGAGCGGCGCCGAGGCTCTCAA 1136
Qy 1448 ACACGGGAATTCACCAAGAGTGTGATGTGTGAGTTTGGGCTGTGCTGTGGAGGT 1507
Db 1137 ACACGGGAATTCACCAAGAGTGTGATGTGTGAGTTTGGGCTGTGCTGTGGAGGT 1196
Qy 1508 CTCTCTCATATGACAGCGGCTCCGTAACCTAAATATGCACTGAAGAGTGTGAGGCGCT 1567
Db 1197 CTCTCTCATATGACAGCGGCTCCGTAACCTAAATATGCACTGAAGAGTGTGAGGCGCT 1256
Qy 1568 GGAGAAAGGGGTACCGCATGTGAACCCCGGAGGGGTGTTCAGAGGCGCCGTGACCTCTCAT 1627
Db 1257 GGAGAAAGGGGTACCGCATGTGAACCCCGGAGGGGTGTTCAGAGGCGCCGTGACCTCTCAT 1316
Qy 1628 GAGCAGCTGCTGGAGAGCAGAGCCGCGCCGCGCACCTTCGCAAACTGGCCGAGAA 1687
Db 1317 GAGCAGCTGCTGGAGAGCAGAGCCGCGCCGCGCACCTTCGCAAACTGGCCGAGAA 1376
Qy 1688 GCTGGCCCGGAGAGCTACAGTCAAGTCAAGTGTGCGCCAGCTCTGCTCAGAGGAGAGCGCA 1747
Db 1377 GCTGGCCCGGAGAGCTACAGTCAAGTCAAGTGTGCGCCAGCTCTGCTCAGAGGAGAGCGCA 1436
Qy 1748 CGGCTCACCTGCGCCGAGAGCAGAGCCCTGACCCCGCTGGGCGCTGGGCGCTGGGCGCC 1807
Db 1437 CGGCTCACCTGCGCCGAGAGCAGAGCCCTGACCCCGCTGGGCGCTGGGCGCTGGGCGCC 1496
Qy 1808 AGAGAGCAGAGAGTGTGAGAGTGTGCGGCTGTGCGGCGCTGACACAGCGCCCAAGAGAGGTCC 1867
Db 1497 AGAGAGCAGAGAGTGTGAGAGTGTGCGGCGCTGTGCGGCGCTGACACAGCGCCCAAGAGAGGTCC 1556
Qy 1868 AGGCGGCAAGTATCTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1927
Db 1557 AGGCGGCAAGTATCTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1616
Qy 1928 GCGGCGCGGTGAGACACCCGACCTGCGAAGATGATGCGCGCTAAAGAGGATTTCTAA 1987
Db 1617 GCGGCGCGGTGAGACACCCGACCTGCGAAGATGATGCGCGCTAAAGAGGATTTCTAA 1676
Qy 1988 GGAAGCTTAAAAA 2000
Db 1677 GGAAAAA 1689

```

```

RESULT 6
AAZ29701
ID AAZ29701 standard; cDNA; 2187 BP.
XX
XX AAZ29701:
XX
XX 22-MAR-2000 (first entry)
XX
XX wild-type human c-Src tyrosine kinase cDNA.
XX
XX Angiogenesis: tyrosine kinase; Src; inhibition; activation; modulate;
XX human; viral expression vector; replication competent; mutant Src;
XX inflammatory disease; arthritis; rheumatoid arthritis; restenosis;
XX diabetic retinopathy; osteoporosis; cancer; ss.
XX Homo sapiens.
XX
XX key location/Qualifiers
XX CDS 134..1486
XX FT /*tag= a
XX FT /product= "Human c-Src tyrosine kinase"
XX FT /note= "Src used to modulate angiogenesis"

```

```

XX
XX W09961590-A1.
XX
XX 02-DEC-1999.
XX
XX 28-MAY-1999; 99WO-US11780.
XX
XX 29-MAY-1998; 98US-0087220.
XX
XX (SCRI) SCRIPPS RES INST.
XX
XX Chersesh DA, Elliceiri B, Schwartzberg PL;
XX
XX WPI: 2000-116335/10.
XX
XX P-PSDB; AAY44448.
XX
XX Using tyrosine kinase Src for modulating angiogenesis in tissues useful
XX in, e.g. treatment of chronic articular rheumatism -
XX
XX Claim 1; Fig 3; 80pp; English.
XX
XX The present sequence is the cDNA, encoding the wild-type human c-Src
XX tyrosine kinase. This sequence encoding the Src protein can be used to
XX modulate angiogenesis. When the Src protein is inactivated, angiogenesis
XX is inhibited, while when it is activated, angiogenesis is potentiated.
XX The modified or variant Src can be used to treat inflammatory diseases
XX like, arthritis, rheumatoid arthritis, diabetic retinopathy, restenosis,
XX osteoporosis and cancer associated disorders.
XX
XX Sequence 2187 BP; 460 A; 651 C; 648 G; 428 T; 0 other;
XX
XX Query Match 27.4%; Score 547.6; DB 21; Length 2187;
XX Best Local Similarity 64.9%; Pred. No. 2,4e-100;
XX Matches 846; Conservative 0; Mismatches 449; Indels 9; Gaps 2;
XX
Qy 395 CTGGGCGCGGCGACCCAGTATACCAATGTGAGACACCCGCCCAAGCAGGGA 454
Db 154 CTGGGCGCGGCGAGTACCAATGTATGCAAGTACCAATTCACGCGCACCTCCGAGCAGA 213
Qy 455 GCTGGCCTTCGCGAGGCGAGCGTGTGACCAATCTGAGAGCGCTGCGAAGAGAGCTG 514
Db 214 CTGGCCTTCGCGAGGAGAGACCGTGTGACCAATTTGTGCGGCTGACCAAGAGACCCCAACTG 273
Qy 515 GTACCGCGTGAAGCACCACACAGTGTGAGAGAGGAGGCTGTGAGCTGGGCGCTGCG 574
Db 274 GTACCAAGGCA---AAACAGAGTGGGCGCTGTGAGGAGTATCCAGCAACTACGTCA 330
Qy 575 GGAAGCGGAGAGCGCTGTGCGCAGACCCAGCTGACAGCTGACGCTGATGCGGTTCCAGGGAA 634
Db 331 GAAAGCGGAGGCGGTGAGAGCGGTACCAAGCTGACGCTGATGCTGTGTTCCAGGCA 390
Qy 635 GATCTGGGCGAGAGAGCTGTGCGAGCTGTGACAGCTGCGGAGATGAGGCTGTCTGTG 694
Db 391 GATCAGACGAGGAGACAGCTGTGAGGCGCTGTGACCGCGGAGACAGGCGCTGTCTGTG 450
Qy 695 GCGGAGAGTCCGCGCGCACCCGCGAGTACGTCTGTGCTGTGAGTGTGCGCGAGCT 754
Db 451 GCGGAGAGAGCACCAATACCCCGAGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 510
Qy 755 CATCTCACTACCGGTGTGACGCGGAGCGGCGCTGACCAATGATGAGGCGGTGCTT 814
Db 511 GAGCACTACCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 570
Qy 815 CTGCAACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 874
Db 571 TGGAACCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630
Qy 875 GCTGTGTGAGACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 934
Db 631 CCTCATTTAAACCAAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 690
Qy 935 GCGCTGTGTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 994

```

[illegible]

PT	XX	/product= "Src"
PM	XX	WO200145751-A1.
PD	XX	28-JUN-2001.
PE	XX	22-DEC-2000; 2000WO-US5396.
PR	XX	22-DEC-1999; 99US-0470881.
PS	XX	29-MAR-2000; 2000US-0538246.
PA	XX	(SCRI) SCRIPPS RES INST.
PB	XX	
PI	XX	Cheresh DA, Elliceiri B, Paul R;
DR	XX	WPI: 2001-417982/44.
DR	XX	P-PSDB; AAB84662.
PT	XX	Modulating vascular permeability in tissues, including inflamed tissue,
PT	XX	tissues associated with stroke, myocardial infarction, by contacting
PT	XX	the tissue with tyrosine kinase protein Src, Yes or their modified
PT	XX	forms -
PS	XX	Disclosure: Fig 3; 133pp; English.
CC	XX	The specification describes a method for modulating vascular
CC	XX	permeability in a tissue suffering from a disease condition. The method
CC	XX	comprises contacting the tissue with a pharmaceutical composition
CC	XX	comprising tyrosine kinase protein Src, Yes or their mixtures or
CC	XX	nucleic acid expressing them. The method is useful for modulating
CC	XX	vascular permeability in tissues, including inflamed tissue, tissues
CC	XX	associated with stroke, myocardial infarction or other blockade of
CC	XX	normal flow, tissues undergoing restenosis, psoriatic, retinal tissue
CC	XX	and similar tissues. Pathologies which may be treated include
CC	XX	trauma to blood vessels, and other systemic pathological events such as
CC	XX	atherosclerosis, diabetic retinopathy, inflammatory disease due to
CC	XX	infection by microbial agents and arthritis. Other diseases which can
CC	XX	be treated include adult respiratory distress syndrome (ARDS), rheumatoid
CC	XX	arthritis, diabetic retinopathy, psoriasis, neovascular glaucoma,
CC	XX	capillary proliferation in atherosclerotic plaques and osteoporosis and
CC	XX	cancer associated disorders such as solid tumours, solid tumour
CC	XX	metastases, angiodiomas and hemangiomas. The present sequence
CC	XX	encodes human Src, and is used in the method of the invention.
SQ	XX	Sequence 2187 BP; 460 A; 651 C; 648 G; 428 T; 0 other:
Query Match	27.4%;	Score 547.6; DB 22; Length 2187;
Best Local Similarity	64.9%;	Pred. No.2.4e-100;
Matches	846; Conservative	0; Mismatches 449; Indels 9; Gaps 2
QY	335	CTGGGCCCCGGGCACCCAGTGTATCATCCAATGCGAGCACACC GCCCCCAACGAGGGA 454
DB	154	CTGGGCATTCCGGTTAAGAATGTAATGCCAACAACTTCACGCGACTGCCGAGCAGGA 213
QY	455	GCTGCGCTTCCGACGAGGGGAGTGATGCATCATCTCTGGAGCGCTCGAAGACAAGCTG 514
DB	214	CCTCGCCCTTCTGCAAAAGAGAGAGCTCTCACCAFTTGTGGCGCTTACCAGAGACCCCACTG 273
QY	515	GTAACCGCTCAAGACACACACACAGTAGAGAGAGGGGCTGTGGCAGCTGGGGCGCTGCG 574
DB	274	GATCAAAAGCCA--AAAACAAGAGTGGGCGGTGAGGAGCATCATCCAGCCAACTACGTCCA 330
QY	575	GGAGGGGAGAGGCCCTCCGCGACAGACCCCAAGTCAAGCTCATGCGTGGTTCCAGGGGA 634
DB	331	GAGAGGGGAGGGCTGAAAGGCGGATACCAAATCTACAGCTTGGTTCCAGCGCA 390
QY	635	GATCTCGGGGACAGGAGGCTGTCCACAGCTGCAGGCTCCCGAGGAGATGGGCTTTCGTG 694
DB	391	GATCAACAGGAGGAGGAGGCTGAGCGGCTTGTACCCGCCGAGACAGGCGCTGTCTGGT 450
QY	695	GGGGAAGTTCGGCGCCGACACCCGGCGAGCTAGCTCTGTGCGTGAAGCTTTGGCCGAGCT 754
DB	451	GGGGGAGGACACCAACTATCCC CGGAGACTACACGCTGTTCGTAAGCTGGAGCGCAAGT 510

Query Match 27.4%; Score 547.6; DB 24; Length 2187;
 Best Local Similarity 64.9%; Pred. No. 2.4e-100;
 Matches 846; Conservative 0; Mismatches 449; Indels 9; Gaps 2;

```

OY 395 CTTGGCCCCCGGCGACCCAGTGTATCACAAATGGAGACACCCGCCCAAGCCAGGGGA 454
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 154 CTTGCCATCCGCGTACAGAAATGTATGGCAAGTACAACTTCCACGGCACTGCCGAGACGA 213
OY 455 GCTGGCCCTTCGGCAAGGGCGACGTGTACCATCTCTGGAGGCTTGGCAAAAGAGCTG 514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 214 CTTGCCCTTTCGCAAGGAGACGTGTACCATCTGTGCGCTTCACCAAGGACCCCACTG 273
OY 515 GTACCGCGTCAAGACACCACTGAGACAGAGGGGCTGTGGAGCTGGGGCGCTCG 574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 GTACAAAGCCA---AAAACAAGGTGGGCGCTGAGGGCATCTCCAGCAACTACGTCCA 330
OY 575 GGAGCGGGAGGCGCTTCCGAGACCCCAAGCTCAGCTCATGGCGTGTTCACAGGGA 634
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 331 GAAGCGGGAGGGCGTGAAGGGCGGTACCAAACTCAGCTCATGGCTTGTTCAGAGGCA 390
OY 635 GATTCGGGGCCAGAGGCTGTCCAGACAGCTCAGCTTCCGAGATGGGCTGTTCGT 694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 391 GATCACAGGGAGGAGGCTGAGCGGCTTCTTACCCGCGGAGACAGGCTGTTCGT 450
OY 695 GCGGGAGTCCGGGCGCCACCCCGGACTACGTCTGTGGTGGAGCTTGGCGCGGAGCT 754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 451 GCGGGAGAGCACCAACTRACCCCGGAGCTACAGCTGTGGCTGAGCTCGAGCGGAGGT 510
OY 755 CATCCACTACCGCGTGTGTGACCGGAGCGGACCTCATCAATCGATGAGGCCGTGTTT 814
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 511 GGAGCAGCTACCGCATCATGTACATGCCAGAGGTGAGATCGACGAGGAGGTGTACTT 570
OY 815 CTGCAACTCATGTGACATGTGTGAGCATTTACAGCAAGACAGAGGCGCTTATCTGACCAA 874
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 571 TGAGAACTCTATGACGTGTGGGAGACTACACTCAGCAGACAGTGTGACTGTATCCGC 630
OY 875 GCTGTGAGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 934
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 631 CTTCTATTAACCAAAAGGTCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 690
OY 935 GGGCTGTGTACTGAACCTGACAGATTGACATTGGGAGACAGATCGAGAGGAGAGT 994
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 691 CGGCTGGGCGCTGAACATGAGAGGAGCTGAGACTGTGACAGCACTGGGAGAGGGAGATT 750
OY 995 TGGAGCTGTCTCTGAGGCTGAGTACCTGGGCAAAAGGTGGCGCTGAAGAAATATCAATG 1054
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 751 CGGAGACGTGATGCTGGGCGATTACCGAGGAGACAAAGTCCCGTCAAGTCAATTAAGA 810
OY 1055 TGAGTGTACAGCCAGGCGCTTCTGGAGAGAGAGGCGGCTCATMGACAGATGCAACAGCA 1114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 811 CGAGGCCAGTCCCAAGGCGCTTCTGGCTGAGAGCCATCATGAGCAACTGGCGGAGAG 870
OY 1115 GAACCTGTGTCTCTCTGGGCGTGTATCTCTGACAG-----GGGCTGTACATTTGCTAT 1168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 871 CAACCTGTGTGACAGCTCTGGGCGTGTATCTGTGAGGAGAGAGGCGGCGCTCATATCTGAC 930
OY 1169 GGAGCAGCTGAGCAAGGGCAACTGTGTGACTTTCTGCGAAGCCGGGGTGCAGCCCTCT 1228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 931 TGAGTACATGTGCCAAGGGGAGGAGCTTGTGACTTCTGGGCTGTAGGGGTCTGAGTGTCT 990
OY 1229 GAACACGCTCAGCTCTGCGAGTTTCTCTGACAGTGGGCGAGGGAGGAGTGAACCTGGA 1288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 991 GGGGAGAGACTGTCTCTCAAGTTCTCGTATGATGTGCGACAGGCGCAATGAAATCTTGA 1050
OY 1289 GAGCAAGAAGCTTGTGACCGCGAGCTGGCGCGCGCAACATCTGTGTCTAGAGAGACT 1348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1051 GGGCAACAATTTTCGATGAGAGAGAGCTGTGGCGGCAATGTGCTGTCTGAGAGCAAA 1110
OY 1349 GGTGGCAAGGTGACGCACTTTGGCTTGGCCCAAGCCGAGCGGAGAGGGGCTTACAGTCAAG 1408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1111 CGTGGCCAAGGTCAAGCGACTTTGTCTTCCCAAGAGGCGTTCACAGCACCGAGGACACGGG 1170

```

```

OY 1409 CCGGCTGCCCTCAAGTGTGACAGGCGCCGAGGCTCTCAAAACAGGAGAGTTCACACAGCAA 1468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1171 CAAGCTGGCACTCAAGTGTGACAGGCGCCCTGAGGCGCTGTGAGAGAGAAATTTCTCACACTA 1230
OY 1469 GTGCGAGTCTGGAGTTTGGGGTGTCTCTCTGGGAGGTCTTTCATATGAGAGCGGCTCC 1528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1231 GTGTGAGCTGTGAGATTTCGGAATCTCTCTGTGGGAAATCTATCTTGTGGCGAGTCC 1290
OY 1529 GTACCTTAATGTCTACAGGAAGAGTGTCCGAGAGCGCTGGAGAGGGGTACCCATGGA 1588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1291 TTATCCAAAGATTTCCCTCAAGAGAGTGTCTCTCGGGTGGAGAGGGCTCAAGATGGA 1350
OY 1589 ACCCGCGAGGCGTGTCCAGGCGCCGCTGACAGCTCTCATGAGACAGCTGTGGAGAGCA 1648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1351 TGCCCCCGAGGCTGCCCGCCCGAGTCTATGAACTATGAAAGACTGTGACACTGCACTGGA 1410
OY 1649 GCCCGCCGCGGCGCCACCTTCCGCAACTGCGCGAGAGCTGG 1692
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1411 GCGCGCATGCGGCGCTCTCTCTATACAGCTCCGAGAGAGAGCTTG 1454

RESULT 9
ABL62918
ID ABL62918 standard; DNA; 2187 BP.
XX AC ABL62918;
XX AC 15-MAY-2002 (first entry)
DT 15-MAY-2002 (first entry)
XX DE Breast cancer related gene sequence SEQ ID NO:1255.
XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumor; cancerous;
KW cytostatic; gene therapy; anti-neoplastic; Wilms' tumour; adenocarcinoma;
KW gene; ds.
XX OS Homo sapiens.
XX PN WO200194629-A2.
XX PD 13-DEC-2001.
XX PD 30-MAY-2001; 2001WO-US10838.
XX PR 05-JUN-2000; 2000US-209473P.
XX PR 05-JUN-2000; 2000US-209531P.
XX PR 18-SEP-2000; 2000US-233133P.
XX PR 18-SEP-2000; 2000US-233617P.
XX PR 20-SEP-2000; 2000US-234009P.
XX PR 20-SEP-2000; 2000US-234034P.
XX PR 20-SEP-2000; 2000US-234052P.
XX PR 20-SEP-2000; 2000US-234509P.
XX PR 22-SEP-2000; 2000US-234567P.
XX PR 22-SEP-2000; 2000US-234923P.
XX PR 25-SEP-2000; 2000US-234924P.
XX PR 25-SEP-2000; 2000US-235077P.
XX PR 25-SEP-2000; 2000US-235082P.
XX PR 25-SEP-2000; 2000US-235134P.
XX PR 25-SEP-2000; 2000US-235280P.
XX PR 26-SEP-2000; 2000US-235637P.
XX PR 26-SEP-2000; 2000US-235638P.
XX PR 27-SEP-2000; 2000US-235711P.
XX PR 27-SEP-2000; 2000US-235720P.
XX PR 27-SEP-2000; 2000US-235740P.
XX PR 27-SEP-2000; 2000US-235863P.
XX PR 28-SEP-2000; 2000US-236028P.
XX PR 28-SEP-2000; 2000US-236032P.
XX PR 28-SEP-2000; 2000US-236033P.
XX PR 28-SEP-2000; 2000US-236034P.
XX PR 28-SEP-2000; 2000US-236109P.
XX PR 28-SEP-2000; 2000US-236111P.
XX PR 29-SEP-2000; 2000US-236842P.
XX PR 29-SEP-2000; 2000US-236891P.

```


PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 (AVAL-) AVALON PHARM.
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppe DR, Weaver Z;
 XX WPI; 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 PS Claim 1; SEQ ID 1255; 44pp; English.
 XX
 XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in AB161664
 CC to AB170110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC adenocarcinoma, carcinoma, kidney, prostate or pancreatic cancer,
 CC infiltrating lobular cancer, clear cell cancer, infiltrating ductal cancer,
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 CC
 SQ Sequence 2187 BP; 460 A; 651 C; 648 G; 428 T; 0 other;
 Query Match 27.4%; Score 547.6; DB 24; Length 2187;
 Best Local Similarity 64.9%; Pred. No. 2,4e-100;
 Matches 846; Conservative 0; Mismatches 449; Indels 9; Gaps 2;

DB 451 GCGGAGAGACCACTACCCCGGAGACTACACGCTGTGCTGAGCGGACAGGT 510
 QY 755 CATCACTACCGGCTGTGACCGGACGCGCACTCACTCAATGATGAGCCGTGCTT 814
 DB 511 GAGACACTACCGCATCATGTACCATGCGACCAAGCTGAGATGACGAGAGGTGCTT 570
 QY 815 CTGCACCTTCATGACATGCTGAGACATTACACCAAGAGAGGCGCTATCTGCACAA 874
 DB 571 TGAAACCTCATGAGCTGTGAGAGCACTACCTCAGACGAGATGAGTCTGTACGCG 630
 QY 875 GCTGTGAGACCAAGGGAAGACAGGAGACCAAGTGGCCGAGAGAGTGGCCAGGCG 934
 DB 631 CTTATTAAACCAAGGTCTATGAGGACACAGTGGCGCCAGAGATGTTTACCCAG 650
 QY 935 GGGCTGTTTACTGACCTGACGATTTGACATTGGGAGCAGACATGCGAGAGGAGATT 994
 DB 691 CGGCTGGGCCCTGACATGAGAGAGCTGAGAGCTGCTGACACCATCGGAGAGGGAGTT 750
 QY 995 TGGAGCTGCTCTGACAGGCTGAGTACCTGGGGCAAAAGGTGGCGTGAAGATATCAAGTG 1054
 DB 751 CGGAGAGCTGATCTGGCGCATTTACGAGGAGAAACAAAGTCCGTCAGATTAAGAA 810
 QY 1055 TGATGTGACACCCAGGCTCTCTGAGACGAGGCGCTGATGAGAAATGGAACGAA 1114
 DB 811 CGAGCCACTGCGAGGCTCTCTGAGAGCTCTGAGAGCTCTGATGAGCACTGGCGCTAG 870
 QY 1115 GAACCTGTGCTGCTCTCTGAGGCTGATCTGACAG-----GGCTGTACATTGTAT 1168
 DB 871 CAACCTGTGCTGCTCTCTGAGGCTGATCTGAGAGAGAGAGGGGCTGTACATGCTGAC 930
 QY 1169 GAGACAGTATGAGAGGCACTGCTGTAACCTTTCTGCGAGACCCGGGCTGAGACCTCTCT 1228
 DB 931 TGAGTATGATGAGAGGAGGAGGCTTGTGACCTACCTGCTGAGGGGCTGCTGCTGCT 990
 QY 1229 GAACACGCTGCTCTCTCTGAGGCTGATCTGACAGGCTGAGGAGGATGAGTACCTGGA 1288
 DB 991 GGGCGGAGAGCTGCTCTCTGAGGCTGATCTGAGAGAGAGAGAGGCTGAGTACCTGGA 1050
 QY 1289 GAGCAGAGAGCTTGTGACAGGAGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1348
 DB 1051 GGGCAGAGAGCTTGTGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1110
 QY 1349 GGTGGCCAGAGTATGAGAGCTTGTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1408
 DB 1111 CGTGGCCAGAGTATGAGAGCTTGTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1170
 QY 1409 CCGGCTGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1468
 DB 1171 CAAGCTGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1230
 QY 1469 GTGCGATGCTGAGAGTGTGGGCTGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1528
 DB 1231 GTCTGAGTGTGAGAGTGTGGAGATCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1290
 QY 1529 GTACCTTAATATGCTACTGAAAGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1588
 DB 1291 TTATCCAGAGATTTCTCTGAAAGAGAGTGTGCTGAGGAGGAGGAGGAGGAGGAGGAG 1350
 QY 1589 ACCCGCGAGAGGCTGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1648
 DB 1351 TGCCCGCGAGAGGCTGCGCCCGCCCACTATGAGAGTATGAGAGAGTGTGAGGAGGAG 1410
 QY 1649 GCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1692
 DB 1411 CGCGCGCATGCGGCGCTCTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1454
 RESULT 10
 AA246489
 ID AA246489 standard; DNA: 2420 BP.
 XX
 XX AA246489;

DT	13-MAR-2000	(first entry)
XX		
DE	PKA substrate; Csk-family protein encoding DNA.	
KW	Protein kinase A; PKA; PKA signaling pathway; phosphorylation; cancer;	
KM	kinase substrate; immunosuppressive disorder; proliferative disease;	
KV	HIV infection; AIDS; immunodeficiency; autoimmune disease; Chk; Lsk; Hyl;	
KW	systemic lupus erythematosus; Csk-family; Matk; Ctk; Bhk; Ntk; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	413..1765
FT		/*tag= a
XX		
PD	02-DEC-1999.	
XX		
PF	27-MAY-1999;	99WO-GB01680.
XX		
PR	27-MAY-1998;	98NO-0002419.
XX		
PR	30-DEC-1998;	98US-0114240.
PA	(LAUR-) LAURAS AS.	
PI	(JONE/) JONES E L.	
PI	Hansson V, Levy FO, Mustelin T, Skalhogg BS, Sundvold V, Tasken K;	
PI	Vang T, Altman A, Munshi A;	
DR	WP1: 2000-086801/07.	
DR	P-PSDB: AA49418.	
XX		
PT	Altering the activity of protein kinase signaling pathways, used for	
PT	treating immunosuppressive disorders, e.g. AIDS, proliferative	
PT	disorders, e.g. cancers or autoimmune diseases	
XX		
XX	Claim 11; Page 87-89; 11pp; English.	
XX		
CC	The invention provides a novel method of altering the activity of the	
CC	protein kinase A (PKA) signaling pathway in a cell that comprises	
CC	altering the extent of phosphorylation of one or more PKA substrates, or	
CC	kinase substrates downstream in the PKA signaling pathway. Pharmaceutical	
CC	compositions containing a nucleic acid molecule that encodes a PKA	
CC	substrate, or fragment, precursor or functionally equivalent variant,	
CC	where the sequence is modified to alter its susceptibility to	
CC	phosphorylation by PKA can be used for treating a disorder exhibiting	
CC	abnormal PKA signaling activity, immunosuppressive disorders or	
CC	proliferative diseases. They can be used for treating e.g. HIV	
CC	infection, AIDS, common variable immunodeficiency or cancers. Conditions	
CC	in which upregulation of the PKA pathway is required, such as autoimmune	
CC	disease, e.g. systemic lupus erythematosus, may also be treated. The	
CC	present sequence represents a DNA sequence encoding a PKA substrate,	
CC	wherein the substrate is in the Csk-family, preferably Csk, Chk, Lsk,	
CC	Hyl, Matk, Ctk, Bhk or Ntk.	
XX		
SQ	Sequence 2420 BP; 471 A; 771 C; 703 G; 475 T; 0 other:	
	Query Match	27.4%; Score 547.6; DB 21; Length 2420;
	Best Local Similarity	64.9%; Freq. No. 2.5e-100;
	Matches 846; Conservative	0; Mismatches 449; Indels 9; Gaps
OY	395 CTGGGCCCCCGGCACCCAGTGTATTCACCAATAATGACAGACACCCTGCCCAAGCAGGGGA	454
Dd		
	433 CTGGCCATCCGCTATCAGAATATGATTTGCCAAATACACTTCCACGGCACTGCCGACACAGA	492
OY	455 GCTGGCCTTCGCGAAGGGCGACGTGGTCACCATCTTGAGAGCCTCGAAGAACAGAGCTG	514
Dd		
	493 CCTGCCTTCGCAAGAGAGACGTCCTCACATTGTGGCGCGTCACCAAGAGACCCCAACTG	552
OY	515 GTACCGGCTCAAGACACACACAGTGGACAGAGAGGGGCTGCTGGCAGCTGGGGCGCTGG	574
Dd		
	553 GTCAAAAGCCA---AAAACAAGGTGGGGCGTGAAGGCAATCAATCCAGCCAACTATGATCA	609

QY	575	GGACGGGAGGCCCTTCCGACAGACCCCAAGCTCATTGCGCTGTTCCACGGGAA	634
Db	610	GAAGCGGAGGGCGCTGAAGCGCGGTACCAAACTAGCTCATGCTTGTCCACGGCAA	669
QY	635	GATCTGGGCCAGGAGCTGTCCAGACCTGCAGACCTCCGAGAGTGGCTGTTCCTGT	694
Db	670	GATCACACGGGAGCGAGCTGAGCGGCTTCTGTACCCCGGAGACAGGCTGTTCCTGT	729
QY	695	CGGGAGTCCGCGCGCCACCCCGCGACTACGTCTGTGGTGAAGCTTGTGGCGGACGT	754
Db	730	GCGGGAGAGCACCAACTACCCCGGAGACTACACGCTGTGGGTGAGTCGACGGCAAGT	789
QY	755	CATCAGTACCGCGCTGTGACACCGCGAGCGACACTCACAATCATGACGCGCTTCTT	814
Db	790	GGACACTTACCGCATATGTATACCATGCGCAGCAAGCTACGATCAGCAGAGAGTACTT	849
QY	815	CTGCACCTCATGAGACATGTTGGAGCTTACAGCAGAGACAAAGGCGCTATTGCAACAA	874
Db	850	TGAAGAACTCATGACGTGTGGAGGACTACACTCAGACGCAAGTAGACTCTGACGC	909
QY	875	GCTGTGAGACCAAGGGGAAACACGGGACCAAGTCGGCGGAGGAGAGCTGGCGCAAGCG	934
Db	910	CCTATTAAACCAAAAGTCAATGGGGACAGTGGCGGCCGCCAGATAGTTCTTACCCGAC	969
QY	935	GGCGTGTCTTACGAACTGCGAGCATTTTGACATTGGGAGACAGATCGAGAGAGAGTT	994
Db	970	CGCGTGGGCCCTGAMCATGAAGGAGCTGAAGCTCTGCAGACCATGGGAAGGGGAGTT	1029
QY	995	TGAGGCTGTCTGTGAGGGTAGTACCTGGGGCAAAAGTGGCGGTGAAGAATATCAAGTG	1054
Db	1030	CGGAGAGCTGATGTGTGGCCATTTACCAAGGGAACAAAGTGGCGCTCAAGTGCATTAAAGAA	1089
QY	1055	TGATGTACACGCCAGGCGCTTCCGTGAGACGAGACGGCGGTGATGACGAAGATGCAACAGA	1114
Db	1090	CGAGCGCACTGCCAGGCGCTTCTGGGTGAAGCTCATGTATGACGCAACTCGCGCATAG	1149
QY	1115	GAACCTGTGCGTCTCCTGGGCGTGATCTGACACAG-----GGGCTGTACATTGTAT	1168
Db	1150	CAACCTGTGAGCTCTCTGGGCGTGATCTGTGAGAGAAAGGCGGCTCTCAATCGTCAAC	1209
QY	1169	GGACGACGTGAGGAGGCAACCTGGTGAATTTTCTCGGAGCCCGGGGTGGAGCCCTCGT	1228
Db	1210	TGAGTACATGCGCCAAAGGAGACCTTCTGCTAGATGTCTCGAGGACATGGAATCTCGGA	1269
QY	1229	GAACACCGCTACGCTCCTGAGTTTCTCTGACAGTGGCGGAGGCGATGAGTACCTGGA	1288
Db	1270	GGGCGGAGACGTGCTCCTCAAGTTCTCGTATGATGTCTCGAGGACATGGAATCTCGGA	1329
QY	1289	GAGCAAAAGCTTGTGACCGCGACCTGGCGGCCCGGCAACATCTGTGTCAGAGACCT	1348
Db	1330	GGGCAAAATTTCTGTGATCGAGACCTGTGCGCCGCAAAATGTGTGTGTGAGGACAA	1389
QY	1349	GGTGGCCAAAGTCAAGCACTTTGGCTCGGCGCAAAAGCCGAGACGGAAGGGGTGACTCAAG	1408
Db	1390	CGTGGCCAAAGTCAAGCACTTTGTGTCTACCAAGAGAGCGCTCCAGCAACCCAGGACAGGG	1449
QY	1409	CGGCGTCCCGTCAAGTGAAGCGCGCCGAGGCTCTCAAAACCGGAAATTTCCACACAA	1468
Db	1450	CAAGCTCCAGTCAAGTGAAGCAGCCCTGAGGCCCTCAGAGAGAAATTTCTCCATTA	1509
QY	1469	GTCGAGTGTGGAGTTTGGGGGTGTGACTGTGGAGAGTCTTCATATGAGACGGGCTCC	1528
Db	1510	GTCGACGTGGAGTTTCGGAATCTTCTCTGGAATATCTACTCTTTTGGCGAGTGC	1569
QY	1529	GTACCCCTAAATGTCACTGAAAGAGTGTGAGAGCCGTGGAGAAAGGGGTACCGCATGA	1588
Db	1570	TTATCCAAAGATTCCCTGAAGGACGCTGTCCTCGGGGTGGAGAAAGGCTTCAAGATGA	1629
QY	1589	ACCGCGGAGGGGTGCCAGGGCCCGGACAGTCTCATAGCAGCTGTGGAGGCGAGA	1648
Db	1630	TGCCCCGAGCGCGCCCGAGCTGTGAAGTATGAAGACTGTGGGACACTGGA	1689

QY 1649 GCCGCCCCGGCCACCTTCCGCAACTGCGCCGAGAGCTGG 1692
DB 1690 CGCGGCATGGCGCCCTCTTCTACAGTCCGAGACGACTTG 1733

RESULT 11
ID ACC50120 standard; cDNA; 2420 BP.
XX
AC ACC50120;
XX
DT 12-JUN-2003 (first entry)
XX
DE Breast cancer associated cDNA sequence SEQ ID NO:87.
XX
KW Human; breast cancer; cytostatic; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200304989-A2.
XX
PD 16-JAN-2003.
XX
PF 21-JUN-2002; 2002WO-US19669.
XX
PR 21-JUN-2001; 2001US-299887P.
PR 27-JUN-2001; 2001US-301572P.
PR 18-JUL-2001; 2001US-306501P.
PR 25-SEP-2001; 2001US-325002P.
PR 05-MAR-2002; 2002US-362585P.
PR 14-MAY-2002; 2002US-380391P.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Lallie JE, Ganavarapu M, Glatc K, Hoersh S, Kamatkar S, Mertens M;
PI Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE, Baer RC;
PI Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;
DR WPI; 2003-210381/20.
XX
P-PSDB; ABR47428.
XX
PT Breast cancer diagnosis or treatment by comparing the level of
PT expression of a marker in a patient sample with that in the control
PT non-breast cancer sample -
XX
PS Claim 1; SEQ ID 87; 128bp; English.
XX
CC The present invention describes a method for assessing whether a patient
CC is afflicted with breast cancer. The method comprises comparing the level
CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
CC ABR47386 to ABR47632) in a patient sample and the normal level of
CC expression of the marker in a control non-breast cancer sample, where a
CC significant increase in the level of expression of the marker in the
CC patient sample and the normal level is an indication that the patient is
CC afflicted with breast cancer. The breast cancer associated sequences
CC from the present invention have cytostatic activities and can be used in
CC gene therapy. The method is useful for diagnosing and treating breast
CC cancer.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
XX
SQ Sequence 2420 BP; 471 A; 771 C; 703 G; 475 T; 0 other;

Query Match 27.4%; Score 547.6; DB 25; Length 2420;
Best Local Similarity 64.9%; Pred. No. 2,5e-100;
Matches 846; Conservative 0; Mismatches 449; Indels 9; Gaps 2;

QY 395 CTGGGCCCCGCGACCCAGTGTATACAAATGCGACACACCGCCCAAGCCAGGGA 454
DB 433 CTGGGCATCCGATACGAATGTATTCGCAAGTACACTTCACGCGACTCCGACGAGA 492
QY 455 GGTGGCCTTCGCAAGGCGACGTGTGACCATCTCTGAGGCTCGCAGAAACAGAGCTG 514

DB 493 CTTGCCCTTGTGCAAAAGAACGTGCTACACATGTGTGCGGTACACCAAGAACCCACTG 552
QY 515 GTACCCGCTCAAGACACACACCAAGTGCAGAGAGGGCTGTGACCTGCGGGCGCTGGC 574
DB 553 GTACAAAGCCA---AAACAAGGTGGGCGCTGAGGGCATATCCACCAACTACGTCCA 609
QY 575 GGAACGGGAGCCCTCTCCGAGACCCCAAGCTCAGCTTATGCGGTGTTCACGGGAA 634
DB 610 GAAGCGGAGGCGGTGAAGCGGGGTACCAACTAGCTCATGCTGTGTTCCAGGCCAA 669
QY 635 GATCTGGGCCAGAGAGCTGTCCAGACGTTCAGACCTCCGAGAGATGGGTCTTCTGGT 694
DB 670 GATTCACAGGGAGAGAGGCTAGCGGCTTGTGTACCCCGAGACAGGCGTGTCTGGT 729
QY 695 GCGGAGTCCGCGCGCCAGCCCGGAGCTACGTCCTGTGCTGAGCTTGGCCGAGAGT 754
DB 730 GCGGAGAGACCAACACTACCCCGAGACTACAGCTGTGCTGAGCTGCGACGCAAGT 789
QY 755 CATCCACTACCGGCTGTGCACCGGAGCGCCACTCACAATCGATGAGGCCGTGTCTT 814
DB 790 GGAAGCACTACCGCATCATGTATCCATGCCACCAAGCTCAGCATCAGCAGAGGATCTT 849
QY 815 CTGCAACTCATGACATGTGTGAGCATTTACAGCAAGAGACAGAGGCCCTATCTCACCA 874
DB 850 TGAAGAACCTCATGACGTGTGAGCACTACCTACAGCGACGAGATGACTGTGTACGCG 909
QY 875 GCTGTGAGACCAAGGAGGAACACGGGACCAAGTGCAGCGAGAGAGGTGGCCAGGCG 934
DB 910 CTCATTTAAACCAAGGTCTATGAGAGGACACAGTGGCGCCAGATGATTTACCCAG 969
QY 935 GGGCTGTTACTGAACTGCAGCATTTGACATTTGGGAGCAAGATCGGAGAGGAGATT 994
DB 970 CGGCTGGGCGCTGACATGAAAGAGAGCTGCTCAGACCATCGGAGGGGAGATT 1029
QY 995 TGAAGCTGTCTGAGAGGTAGTACCTGTGGGCAAAAGTGGCGGTGAAGATATCAAGTG 1054
DB 1030 CGGAGACGTATGCTGGCGATTACGAGGAAACAAAGTCCGCTCAAGTGCATTAAAGAA 1089
QY 1055 TGAATGACAGCCAGCCGCTTCTCGAGAGAGCGCCGCTATGACGAAGATGCAACAGA 1114
DB 1090 CGAGCCGACAGCCAGGCGCTTCTCGAGTGAAGCTCAGTATGACGACACTCGGCGATAG 1149
QY 1115 GAACCTGTGCTGCTCTGGGCGTGTATCTGCACAG-----GGCTGTACATTTGTCAT 1168
DB 1150 CAACCTGTGCTGCTCTGGGCGTGTATCTGCAGAGAGAGGCGGCTCTACATCTGCAC 1209
QY 1169 GAGACAGTGTGAGCAAGGGAACCTGTGAGACTTCTGCGGACCCGCGGTGAGCCCTGT 1228
DB 1210 TGAATACATGCGCAAGGGAGCCCTGTGTGACTACTGCTGAGGGGTGCTAGTGTCT 1269
QY 1229 GAACACCGCTCAGCTCTGCAAGTTTCTGTGACAGTGGCGAGGAGTGTGAGTACCTGGA 1288
DB 1270 GGGCGGAGACTGTCTCTCAAGTTCTGTGATGTGTGTGGAGGCCATGGAATACCTGGA 1329
QY 1289 GAGCAGAGACTGTGTGACCGGACCTGCGCCCGCCGACATCTGTCTCAGAGACCT 1348
DB 1330 GGGCAACAAATTTGTGTGATGAGACCTGTGCGCCGCAAAATGTGTGTCTCAGGCAAA 1389
QY 1349 GGTGGCCAAAGTGTGAGCACTTGTGGCCGTGGCAAAAGCGAGCGGAGGCGGTACACTGA 1408
DB 1390 GGTGGCCAAAGTGTGAGCACTTGTGTGTCTCAGCAAGAGGCGCTCAGCAGCAGGAGG 1449
QY 1409 CCGGCTGCCCTGAAGTGTGAGCGGCGCGAGGCTCTCAACACGAGGAAGTTCCACGACAA 1468
DB 1450 CAAGCTGCCAGTGTGAGTGTGAGCGAGCGCCGTGAGGCGCTGAGAGAAAGAAATTTCCACTAA 1509
QY 1469 GTGAGATGTGTGAGTTTGGGCTGTCTGTGTGAGAGTCTTGTATATGACAGGCTCTC 1528
DB 1510 GTCTGAGTGTGTGAGTTTGGGAAATCTTCTGTGTGAGAAATCTCTTGTGTGGAGAGTCC 1569
QY 1529 GTACCCTAATAATGTCACTGAAAGAGTGTGTGAGGCGGTGTGAGAGGGGTATCCGATGGA 1588

QY 1529 GTACCTTAATGTCACCTGGAAGAGCTGTGCGAGCCGTGAGAGGGGTACCCATGCA 1588
 Db 1570 TTATTCAGAAATTCCTCCGTAAGGAGCTGTGCTCTCGGGTGGAGAGGGCTCAAGATGGA 1629
 QY 1589 ACCCCCGAGGGCTGTGTCAGGCGCCGTCGACGCTCTCATGAGCTGTGCGAGGAGCA 1648
 Db 1630 TGCCCCCGAGGGCTGTGTCAGGCGCCGTCGACGCTCTCATGAGCTGTGCGAGGAGCA 1689
 QY 1649 GCG 1692
 Db 1690 CG 1733

RESULT 13

ABST73326 standard; DNA: 2442 BP.
 ABST73326:
 04-DEC-2002 (first entry)
 DNA encoding human c-src isoform.

Chromosome aberration; oncogenic fusion protein; cancer; oncogene;
 proliferative disease; cellular protein isoform; heat shock protein 90;
 HSP-90; rheumatoid arthritis; cancer; hematopoietic disorder;
 T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
 acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
 acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
 papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
 rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.

Homo sapiens.
 MO200269900-A2.
 12-SEP-2002.
 01-MAR-2002; 2002MO-US06518.
 01-MAR-2001; 2001US-272751P.
 (CONF-) CONFORMA THERAPEUTICS CORP.

Fritz LC, Burrows FJ;
 WPI: 2002-698710/75.
 P-PSDB; ABG95117.

Treating genetically-defined disease associated with chromosomal
 aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
 diseases, involves administering an inhibitor of heat shock protein 90

Disclosure: Page 313-314; 389pp; English.

The invention describes a method of treating genetically-defined disease
 associated with chromosomal aberrations yielding oncogenic fusion
 proteins (I), treating cancerous cells containing (I) in a heterogeneous
 cell population, treating proliferative diseases associated with mutant
 protein or cellular protein isoforms (II) dependent on heat shock
 protein (HSP)-90, or selectively treating cells expressing (II)
 involving administering HSP90-inhibitor. The method is useful for
 treating genetically-defined disease with chromosomal aberration yielding
 oncogenic fusion protein, treating cancerous cells containing fusion
 protein in heterogeneous cell population, treating proliferative disease
 (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
 cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
 p53), or selectively treating cells expressing mutant protein or cellular
 protein isoform in a patient heterozygous for (II). The method is useful
 for treating a disease e.g. hematopoietic disorder such as T or B cell
 lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CMML.

CC or a disease characterised by a solid tumour such as papillary thyroid
 carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
 synovial sarcoma. The method is also useful for treating viral
 infections. This sequence encodes a human oncogenic protein.

Sequence 2442 BP; 474 A; 779 C; 708 G; 481 T; 0 other:

Query Match 26.4%; Score 527.8; DB 24; Length 2442;
 Best Local Similarity 64.9%; Pred. No. 2.3e-96;
 Matches 847; Conservative 0; Mismatches 447; Indels 11; Gaps 4;

395 CTGGGCCCCGGGCAACCAAGTATACCAAAATGGAGACACACCCGCCCAAGCCAGGGA 454
 Db 459 CTGGGCAATCCGTCACAAATGATATGGCAATACCTTCACGCACTGCCGAGGCA 518
 QY 455 GCTGGCTTCCGCAAGGCGCAGTGTACCATCTCGAGGCGTGGAGAAAGAGCTG 514
 Db 519 CTTGCCCTTTCGCAAGAGACGTCTCACCATTGTGGCGTCCACAAAGAACCCAACTG 578
 QY 515 GTACCGGCTCAAGCACACACCAAGTGAAGAGAGGGCTGTGACCTGGGCGCTGCG 574
 Db 579 GTACAAAGCCA---AAACAAAGTGGGCGCGTGAAGGCGCATCATCCACGCAATCGTCA 635
 QY 575 GGAGCGGAGGCGCTCTCCGACACCCCAAGCTTACCTCATGCGCTG-CTTCACGGGA 633
 Db 636 GAAACGGGAGGCGCTTGAAGCGGGGTACCAAACTCAGCTCATGCGCTGAGTTCACGGCA 695
 QY 634 AGATCTGGGCGCAAGAGAGCTGTGACAGAGCTGCGAGCTCCCGAGATGGGCTGTCTGG 693
 Db 696 AGATCTACAGGAGAGAGCTGTGAGCGGCTTGTATACCCCGGAGACAGGCTGTCTGG 755
 QY 694 TCGCGAGTCCGGCGCGCCACCGCGCACTACGTCTGTGCTGTGGCGCGAGC 753
 Db 756 TCGCGAGAGACACCAACTACCCGAGAGATACAGCTGTGCTGTGAGCTGGAGCGCAAG 815
 QY 754 TCATCCACTACCGGTCTGTGACCGGAGCCACCTCACAATGATGAGCGGTGTCT 813
 Db 816 TCGACACTACCGGTATGTATCATCTCCAGCAAGCTCAGCATGAGAGAGGTGTACT 875
 QY 814 TCTGCAACTCATAGAGATGTGTGAGAGCATTTGACATTTGGAGAGCAAGATCGGAGAGAGT 933
 Db 876 TTGAGAACCTCATGCACTGTGTGTGAGAGCATCTACACCTCAGACGAGATGAGTGTGTCG 935
 QY 874 AGCTGTGAGACCAAGCGGAACACGAGCAAACTGTGCGCGAGAGAGAGCTGGCCAGG 933
 Db 936 GCTCATTTAAACCAAGGTATGAGGCGCACAGTGGCGGCCAGATGATGTATACGCCA 995
 QY 934 CGGGCTGTACTGAACTGTGACATTTGATTTGGAGAGCAAGATCGGAGAGAGAGT 993
 Db 996 GCGGCTGGGCTGTGAACTGTGAGAGGCTGAGAGCTGTGACAGACCATCGGAGAGAGT 1055
 QY 994 TTGAGAGCTGTCTGCAAGGAGTGTGAGTGTGAGGCAAAAGTGTGAGGCTGTGAAATATCAAGT 1053
 Db 1056 TCGGAGACGTGATGTGTGGGAGTATACGAGGAAACAAAGTGTGCGCTCAATGATTAAGA 1115
 QY 1054 GTGATGTGACAGCCAGGCGCTTCTGAGAGAGAGCGGCTCATGAGAGATCAACAGC 1113
 Db 1116 ACCAGCGCACTGCGCAGGCGCTTCTGAGAGAGAGCGCTCATGAGAGAGAGAGT 1175
 QY 1114 AGAAGCTGTGCTGTCTGTGGGCTGTGATCTGACAGCAG-----GGGCTTACATTTGTCA 1167
 Db 1176 GCAACCTGTGTGAGCTGTGCGGTGTGATCTGAGAGAGAGGCGGCTTCAATCGTCA 1235
 QY 1168 TGAGACAGTGAGT 1227
 Db 1236 CTGAGTACATGAGT 1295
 QY 1228 TGAACACCGCTACAGTCTGTGACATTTCTGTGACAGTGGCGAGAGAGAGAGAGAGT 1287
 Db 1296 TGCGGAGAGAGTGTCTGTCAAGTGTCTGATGATGTGAGAGAGAGAGAGAGAGT 1355
 QY 1288 AGAGCAAGAGAGTGTGTGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1347

```
Db 1356 AGGCGACAAATTTCTGCATCGAGACGTGCTGCCCGCCGAAATGTCTGTCTGAGAGACA 1415
QY 1348 TGGTGGCCAGGTCAGCGACTTTGGCTGGCCAAAGCGAGAGGGGCTAGACTCAA 1407
Db 1416 ACGTGGCCAGGTCAGCGACTTTGGCTGGCCAAAGCGAGAGCG-TCACACCCGAGAGACGG 1474
QY 1408 GCCGGCTCCCGTCAGTGAAGCGGCCCGGAGGCTCTCAACACGGAAGTTCAACAGCA 1467
Db 1475 GCAGCTGCAGTCAAGTGAAGACAGCCCTGAGAGAGAGAAATCTCTCACACIA 1534
QY 1468 AGTCGATGTCTGAGATTGGGGTGTCTCTGAGAGGCTTCTCAATATGAGAGCGGCTC 1527
Db 1535 AGCTGACGTGTGGAGTTCTCGGAATCTCTCTGGGAATCTCTTGGCGAGTGC 1594
QY 1528 CGTACCTTAATATGCTCAAGAGGTGTGGAGCGCGGTGAGAGAGGGTACCCCATGG 1587
Db 1595 CTTATTCAGAGATTCCTCGAAGAGAGCTGCTCCGTGGGTGAGAGGGCTACAAAGATGG 1654
QY 1588 AACCCCGGAGGCTGTCTCAGGCCCCGTCACAGTCTCATGACAGTCTGGAGAGCAG 1647
Db 1655 ATGCCCCGAGCGCTGCCCGCCGCGAGTCTATGAGTCAATGAGAACTGCTGGACCTGG 1714
QY 1648 AGCCCGCCCGCCGCGCACCCCTTCGCCAAACTGCGCGAGAGACTGG 1692
Db 1715 ACGCGCGCATGCGGCCCTCTCTTCTACAGCTCCGAGAGACGCTTG 1759

RESULT 14
ABK88792 standard; DNA; 16389 BP.
XX
AC ABK88792;
XX
XX 21-OCT-2002 (first entry)
XX
DE Gene encoding human kinase related to tyrosine kinase family.
XX
KW Human; kinase; tyrosine kinase; kinase mediated disorder;
KW haematopoietic disorder; cell signal transduction disorder;
KW cancer; haemostatic; cytosstatic; chromosome 19; gene; ds; SNP;
KW single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
XX
XX KEY Location/Qualifiers
XX FH 3000..14181
XX FT /tag= a
XX FT /product= "Kinase"
XX FT /number= 3085
XX FT /tag= b
XX FT /number= 1
XX FT 3086..8330
XX FT /tag= c
XX FT /number= 1
XX FT 8331..8550
XX FT /tag= d
XX FT /number= 2
XX FT 8551..9144
XX FT /tag= e
XX FT /number= 2
XX FT 9145..9238
XX FT /tag= f
XX FT /number= 3
XX FT 9239..10692
XX FT /tag= g
XX FT /number= 3
XX FT 10693..10758
XX FT /tag= h
XX FT /number= 4
XX FT 10759..12566
XX FT /tag= i
XX FT /number= 4
XX FT 12567..12666
XX FT exon

FT FT /tag= j
FT FT /number= 5
FT FT 12667..12746
FT FT /tag= k
FT FT /number= 5
FT FT 12747..12831
FT FT /tag= l
FT FT /number= 6
FT FT 12832..12912
FT FT /tag= m
FT FT /number= 6
FT FT 12913..12986
FT FT /tag= n
FT FT /number= 7
FT FT 12987..13176
FT FT /tag= o
FT FT /number= 7
FT FT 13177..13372
FT FT /tag= p
FT FT /number= 8
FT FT 13373..13768
FT FT /tag= q
FT FT /number= 9
FT FT 13769..13855
FT FT /tag= r
FT FT /number= 9
FT FT 13856..13941
FT FT /tag= s
FT FT /number= 9
FT FT 13942..14178
FT FT /tag= t
FT FT /number= 10
FT FT (replace 1130, A)
FT FT /tag= u
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT (replace 4245, A)
FT FT /tag= v
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT (replace 4248, T)
FT FT /tag= w
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT (replace 4250, A)
FT FT /tag= x
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT (replace 4255, A)
FT FT /tag= y
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT (replace 4258, T)
FT FT /tag= z
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT (replace 4260..4261, T)
FT FT /tag= aa
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT (replace 4266..4267, T)
FT FT /tag= ab
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT (replace 4269..4270, T)
FT FT /tag= ac
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT (replace 4273, T)
FT FT /tag= ad
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT (replace 4275, A)
FT FT /tag= ae
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT (replace 4401, T)
FT FT /tag= af
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT (replace 4551..4552, C)
FT FT /tag= ag
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT (replace 4739, G)
FT FT /tag= ah
```

FT variation /standard_name="Single nucleotide polymorphism (SNP)"
 FT (replace 6309, A)
 FT /tag= a1
 FT /standard_name="Single nucleotide polymorphism (SNP)"
 FT (replace 8753..8754, T)
 FT /tag= a1
 FT /standard_name="Single nucleotide polymorphism (SNP)"
 FT (replace 10797, T)
 FT /tag= ak
 FT /standard_name="Single nucleotide polymorphism (SNP)"
 FT (replace 10806, G)
 FT /tag= al
 FT /standard_name="Single nucleotide polymorphism (SNP)"
 FT (replace 11199, T)
 FT /tag= am
 FT /standard_name="Single nucleotide polymorphism (SNP)"
 FT (replace 11201, T)
 FT /tag= an
 FT /standard_name="Single nucleotide polymorphism (SNP)"
 FT (replace 11319, C)
 FT /tag= ao
 FT /standard_name="Single nucleotide polymorphism (SNP)"
 FT (replace 12238, C)
 FT /tag= ap
 FT /standard_name="Single nucleotide polymorphism (SNP)"
 FT (replace 14739, T)
 FT /tag= aq
 FT /standard_name="Single nucleotide polymorphism (SNP)"
 FT (replace 15375, C)
 FT /tag= ar
 FT /standard_name="Single nucleotide polymorphism (SNP)"
 PN WO200252018-A2.
 PD 04-JUL-2002.
 XX 19-DEC-2001; 2001WO-US48546.
 XX 21-DEC-2000; 2000US-0741154.
 PA (PEKE) PE CORP NY.
 PI Beasley EM, Shao W, Ketchum K, Di Francesco V;
 DR WPI; 2002-583568/62.
 DR P-PSDB; AA099575.
 XX
 XX New Isolated human kinase proteins and genes, useful in developing
 PT drugs, as well as for diagnosing, preventing or treating disorders
 PT associated with defective cell signal transduction, e.g. cancer or
 PT hematopoietic disorders
 XX
 PS Claim 4; Fig 3A-F; 78pp; English.
 XX
 CC The present invention relates to the isolation of a novel human kinase
 CC related to the tyrosine kinase family, and polynucleotide sequences
 CC encoding it. The gene encoding the kinase of the invention maps to
 CC chromosome 19. The kinase may be used for identifying a modulator of
 CC the kinase, an agent that binds to the kinase, or for identifying
 CC other members of the family. The kinase may also be used to raise
 CC antibodies which may be used in immunoassays or drug screening assays.
 CC The sequences of the invention may be used for treating a disease or
 CC condition mediated by a human kinase such as hematopoietic disorders
 CC and other disorders associated with defective cell signal transduction
 CC (e.g. cancer). The polynucleotide sequences encoding the kinase are
 CC useful for isolating and purifying the kinase, and as probes or primers.
 CC The present sequence encodes the human kinase of the invention.
 XX
 SO Sequence 16389 BP; 3511 A; 4901 C; 4515 G; 3462 T; 0 other;

Query Match 22.8%; Score 455; DB 24; Length 16389;
 Best Local Similarity 100.0%; Pred No. 11e-81;
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1541 GTCACGAAAGAGCTGCGAGGCGGTGAGAGAGGGGTACCGCATGGAACCCCGAGGG 1600
 |||||||
 Db 13941 GTCACGAAAGAGGTGTGCGAGGCGGTGAGAGAGGGGTACCGCATGGAACCCCGAGGG 14000
 |||||||
 QY 1601 CTGTCCAGGCGCGGTGCGAGGCTGCTATGACGAGCTGTGGAGGAGAGCCCGCCGCG 1660
 |||||||
 Db 14001 CTGTCCAGGCGCGGTGCGAGGCTGCTATGACGAGCTGTGGAGGAGAGCCCGCCGCG 14060
 |||||||
 QY 1661 GCCACCGTTCGCCAAGCTGCGGAGAGAGCTGCGGAGCTACGCGAGTGCAGGTGCCCC 1720
 |||||||
 Db 14061 GCCACCGTTCGCCAAGCTGCGGAGAGCTGCGGAGCTACGCGAGTGCAGGTGCCCC 14120
 |||||||
 QY 1721 AGCTTCGCTCAGGCGAGGAGCGGAGGCTTCCACTTCGCGCCGAGCGAGCCCTG 1780
 |||||||
 Db 14121 AGCTTCGCTCAGGCGAGGAGCGGAGGCTTCCACTTCGCGCCGAGCGAGCCCTG 14180
 |||||||
 QY 1781 ACCCCAGCGGTGGGGCCCTTGGCCGAGAGAGCGAGAGTGTGAGAGTGGCGGTGG 1840
 |||||||
 Db 14181 ACCCCAGCGGTGGGGCCCTTGGCCGAGAGAGCGAGAGTGTGAGAGTGGCGGTGG 14240
 |||||||
 QY 1841 GGCACGTGACAGGCGCCAGAGAGGTCAGGCGGCGCAAGTCACTCTGTGTGCCACAGC 1900
 |||||||
 Db 14241 GGCACGTGACAGGCGCCAGAGAGGTCAGGCGGCGCAAGTCACTCTGTGTGCCACAGC 14300
 |||||||
 QY 1901 AGGGCGTGGCCGCTAGGGGCTGTGGGCGCCGCTGTGACACCCAGACCTGCGAAGA 1960
 |||||||
 Db 14301 AGGGCGTGGCCGCTAGGGGCTGTGGGCGCCGCTGTGACACCCAGACCTGCGAAGA 14360
 |||||||
 QY 1961 TGATCGCCCGATTAAGACGATTCCTTAAGACTCTA 1995
 |||||||
 Db 14361 TGATCGCCCGATTAAGACGATTCCTTAAGACTCTA 14395
 |||||||

RESULT 15
 ABL11799
 ID ABL11799 standard; cDNA; 2591 BP.
 XX
 AC ABL11799;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 29879.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 OS
 OS Drosophila melanogaster.
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 DR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EM;
 DR WPI; 2001-656860/75.
 DR P-PSDB; ABB67696.
 XX
 PT New Isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions
 XX
 PS Claim 1; SEQ ID NO 29879; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 15:09:19 ; Search time 5073 Seconds
9581.895 Million cell updates/sec

Title: US-09-977-260-1

Perfect score: 2000
Sequence: 1 ctgcctcaagtgtgtgcagc.....attctaaggaacttaaaaa 2000

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1071.8	53.6	2046	11 AK013606
2	926.2	46.3	1035	9 AL580639 Mus muscu
3	924	46.2	1040	9 AL558929 AL558929
4	909.6	45.5	1032	9 AL558805 AL558805

C 5	867	43.4	906	10 BG744770	BG744770 602722891
C 6	857.2	42.9	1012	9 AL580543	AL580543 AL580543
C 7	845.4	42.3	1043	12 B0054024	B0054024 AGENCOURT
C 8	830	41.5	949	9 AL567349	AL567349 AL567349
C 9	811.4	40.6	1102	12 BM561134	BM561134 AGENCOURT
C 10	799	40.0	967	9 AL538511	AL538511 AL538511
C 11	795.4	39.8	862	12 B1768288	B1768288 603056458
C 12	762	38.1	912	12 B1819535	B1819535 603036746
C 13	737.2	36.9	1014	12 BM925586	BM925586 AGENCOURT
C 14	707.6	35.4	733	13 B0617908	B0617908 UT-H-DPO-
C 15	707	35.4	986	13 BX369553	BX369553 BX369553
C 16	693.4	34.7	842	12 B1910376	B1910376 603068035
C 17	693	34.6	712	12 BM681434	BM681434 UT-E-EJO-
C 18	691.4	34.6	1129	12 BM450510	BM450510 AGENCOURT
C 19	680.6	34.0	936	13 BU184285	BU184285 AGENCOURT
C 20	679.6	34.0	822	10 BG358801	BG358801 602458975
C 21	677.2	33.9	772	10 BG758878	BG758878 602713275
C 22	674.4	33.7	805	10 BE781641	BE781641 601470504
C 23	671.8	33.6	1119	12 BM925766	BM925766 AGENCOURT
C 24	661	33.1	708	12 BM715548	BM715548 UT-E-EJO-
C 25	660.4	33.0	702	12 BM674009	BM674009 UT-E-EJO-
C 26	641.2	32.1	732	12 B1769001	B1769001 603058156
C 27	640.4	32.0	837	12 B1912704	B1912704 603176356
C 28	626.4	31.3	651	13 B0582182	B0582182 1113910.x
C 29	624.6	31.2	741	12 BM719266	BM719266 UT-E-EJO-
C 30	621.4	31.1	729	9 AU135553	AU135553 AU135553
C 31	621.2	31.1	767	12 B1911599	B1911599 603064181
C 32	608	30.4	953	12 B1488576	B1488576 603021104
C 33	602.4	30.1	729	13 B0070968	B0070968 Im28C03.x
C 34	602.2	30.1	729	9 AV702542	AV702542 AV702542
C 35	599.6	30.0	931	12 B1549778	B1549778 603194514
C 36	598	29.9	735	10 BE779216	BE779216 601464660
C 37	597.6	29.9	847	10 BG396388	BG396388 602459315
C 38	595	29.8	607	10 BE257861	BE257861 601109450
C 39	593	29.6	593	13 B0073359	B0073359 Im34905.x
C 40	592.2	29.6	631	10 BE254257	BE254257 601114671
C 41	589	29.4	613	10 BG396034	BG396034 602458772
C 42	587.6	29.4	937	13 B0438002	B0438002 AGENCOURT
C 43	577.8	28.9	1138	10 BE791829	BE791829 601586106
C 44	576.2	28.8	816	10 BE783096	BE783096 601470686
C 45	571	28.5	594	13 B0077460	B0077460 Im36906.x

ALIGNMENTS

RESULT 1
AK013606
LOCUS
DEFINITION
Mus musculus adult male hippocampus cDNA, RIKEN full-length
enriched library, clone:2000029A15 product:megekaryocyte-associated
tyrosine kinase, full insert sequence.
ACCESSION
AK013606
VERSION
AK013606.1 GI:12851034
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL
MEDLINE
20499374
PUBMED
11042159


```
Db 508 CGTGACGTATCATCACTACCGTGTGTTTGCATGAGATGGGACCTTCACCATCGATGAGAGCC 567
Oy 807 GTGTTCTTCTGCAACCTCATGAGATGSGAGAGCTTACACCAAGAGAGAGGCCCTATC 866
Db 568 GTGTTCTTCTGCAACCTCATGAGATGSGAGAGCTTACACCAAGAGAGAGGCCCTATC 627
Oy 867 TGCACCAAGCTGTGTAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 926
Db 628 TGCACCAAGCTGTGTAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 687
Oy 927 GCCAGAGGCGGCGTGTACTACCACTGACGATTTGACATTTGGAGACACAGATCGAGAG 986
Db 688 GCCAGAGGCGGCGTGTACTACCACTGACGATTTGACATTTGGAGAGAGAGATTTGAGAGAT 747
Oy 987 GGAGGTTTGGAGGCTGCTGACAGGAGTACCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1046
Db 748 GGGAGTTTGGAGGCGTGTACTACAGAGGATGATACCTGGGAGAGAGAGAGAGAGAGAGAGAT 807
Oy 1047 ATCAAGTGTATGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 1106
Db 808 ATCAAGTGTATGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 867
Oy 1107 CAACACAGAAACCTGTGTCTGCTGCGGTGATCTGACACAGAGGCGCTGATCTGATCTG 1166
Db 868 CAGACAGAGAACTGTGTCTGCTGCGGTGATCTGACACAGAGGCGCTGATCTGATCTG 927
Oy 1167 ATGGAGCAGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1226
Db 928 ATGGAGCAGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 987
Oy 1227 GTGAGACAGGCTGTCAAGCTCTGCAAGTTTCTGTGACAGTGGCGGAGAGAGAGAGAGATCTG 1286
Db 988 GTGAGACAGGCTGTCAAGCTCTGCAAGTTTCTGTGACAGTGGCGGAGAGAGAGAGATCTG 1047
Oy 1287 GAGAGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1346
Db 1048 GAGAGCAAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1107
Oy 1347 CTGAGGCGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1406
Db 1108 TTGAGGCGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1167
Oy 1407 AGCGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1466
Db 1168 AGCGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1227
Oy 1467 AAGTCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1526
Db 1228 AAGTCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1287
Oy 1527 CCGTACCTTAATATGTCTGAAAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1586
Db 1288 CCGTACCTTAATATGTCTGAAAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1347
Oy 1587 GAACCCCGGAGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1646
Db 1348 GAGCGCGGAGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1407
Oy 1647 GAGCGCGGAGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1706
Db 1408 GAGCGCGGAGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1467
Oy 1707 AGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1766
Db 1468 AGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1527
Oy 1767 AGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1785
Db 1528 AGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1546
```

RESULT 2
AL580639/c

```
LOCUS AL580639 1035 bp mRNA linear EST 01-JUN-2003
DEFINITION Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
ACCESSION AL580639
VERSION AL580639.2 GI:31318906
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1035)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 16, 2001 this sequence version replaced gi.12946853.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9238.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ007AF03NP1cluster-9238.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DJ007AF03NP1.
FEATURES
source
1..1035
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ007YK05"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 166 a 327 c 304 g 213 t
ORIGIN
Query Match 46.3% Score 926.2; DB 9; Length 1035;
Best Local Similarity 97.3%; Fred. No. 5.9e-173;
Matches 960; Conservative 10; Mismatches 14; Indels 3; Gaps 3;
Oy 936 GGCTGTGTACTGAACTGACCTGACATTTGTGGAGACAGATCGAGAGAGAGTTT 995
Db 1035 GGCTGTGTACTGAACTGACCTGACATTTGTGGAGACAGATCGAGAGAGAGTTT 976
Oy 996 GGAGCTGTCTGTGAGAGGAGTACTGCTGGGCAAAAAGTGGCCGTGAAGATATCAAGTGT 1055
Db 975 GGAGCTGTCTGTGAGAGGAGTACTGCTGGGCAAAAAGTGGCCGTGAAGATATCAAGTGT 916
Oy 1056 GATGTGAGAGCCGAGGCTTCTGTGAGACAGAGCGCCGTCATGAGAGAGATGACACAGAG 1115
Db 915 GATGTGAGAGCCGAGGCTTCTGTGAGAGAGCGCCGTCATGAGAGAGATGACACAGAG 856
Oy 1116 AACCTGGGTCTTCCTGGGGGTGATCTGTGACCAAGAGGGCTGTACATTTGATGAGAGAC 1175
Db 855 AACCTGGGTCTTCCTGGGGGTGATCTGTGACCAAGAGGGCTGTACATTTGATGAGAGAC 796
Oy 1176 GTGAGCAAGAGGCAACTGTGTGTAATTTTGTGGGAGCCCGGGGTGAGAGCCCTGTGAACAC 1235
Db 795 GTGAGCAAGAGGCAACTGTGTGTAATTTTGTGGGAGCCCGGGGTGAGAGCCCTGTGAACAC 736
Oy 1236 GCTCAGCTCCTGTGCAATTTTCTGTGACAGTGGCCGAGAGGAGATGAGTACTGTGAGAGCAAG 1295
Db 735 GCTCAGCTCCTGTGCAATTTTCTGTGACAGTGGCCGAGAGGAGATGAGTACTGTGAGAGCAAG 676
Oy 1296 AAGCTTGTGACCCGAGACTGGCGCCGCAACATCTGTGTCTGAGAGAGACTGTGGGCC 1355
```

```

Db      675  AAGCTGTGACGACGACGCTGGCCGCCGCAACATCTGTTCTCAGAGAACCTGTGGCC 616
QY      1356  AAGGTACGACTTTTGGCCCTGAGCAAGCGAGAGAGGGGTAGACTCAAGCCGGCTG 1415
Db      615  AAGTTCAGGACACTTTGGCTGGCCCAAGCCAGAGAGGGGTAGACTCAAGCCGGCTG 556
QY      1416  CCGCTCAAGTGAAGCGGCGCCGAGGCTCTCAACACGCGAAGTTTCACCAAGTGGAT 1475
Db      555  CCGCTCAAGTGAAGCGGCGCCGAGGCTCTCAACACGCGAAGTTTCACCAAGTGGAT 496
QY      1476  GTCGTGAGTTTGGGAGTGTCTGAGAGTCTTCTCATATGAGAGGGCTCCGTACCT 1535
Db      495  GTCTGAGTTTGGGAGTGTCTGAGAGTCTTCTCATATGAGAGGGCTCCGTACCT 436
QY      1536  AAATGTCACTGAAAGAGTGTGAGAGCGGTGAGAGAGGGGTACCGCATGAAACCCCC 1595
Db      435  AAATGTCACTGAAAGAGTGTGAGAGCGGTGAGAGAGGGGTACCGCATGAAACCCCC 376
QY      1596  GAGGCTGTCCAGAGCGCCGTCGACGTCCTCATGAGCAGCTGTGGAGAGAGCCGCC 1655
Db      375  GAGGCTGTCCAGAGCGCCGTCGACGTCCTCATGAGCAGCTGTGGAGAGAGCCGCC 316
QY      1656  CGCGGCGCAACCTTCGCAAACTGGCGGAGAAAGCTGGCCGGAGCTACGCACTGAGT 1715
Db      315  CGCGGCGCAACCTTCGCAAACTGGCGGAGAAAGCTGGCCGGAG-CTACGCACTGAGT 257
QY      1716  GCCCGACCTCCGCTCAGAGGAGAGAGCGGCGGCTCAGCTCGGCCCGCAAGCCAGAG 1775
Db      256  GCCCGACCTCCGCTCAGAGGAGAGAGCGGCGGCTCAGCTCGGCCCGCAAGCCAGAG 197
QY      1776  CCTGACCCCAAGCCGCTGGGGCCCTTGGCCCAAGAGACCAGAGAGTGGAGAGTGGG 1835
Db      196  CCTGACCCCAAGCCGCTGGGGCCCTTGGCCCAAGAGACCAGAGAGTGGAGAGTGGG 138
QY      1836  GTGGGGGCACTGACAGAGGCGCAAGGAGGGTCCAGGGGGGCAAGTCAATCCCTGGTGG 1895
Db      137  GTGGGGGCACTGACAGAGGCGCAAGGAGGGTCCAGGGGAGCAATCAATCCCTGGGCCC 78
QY      1896  ACAGAGGGGCTGGCCACGTAAGGGG 1922
Db      77  ACAGCA-GGGCTGGCCACGTAAGGGB 52

RESULT 3
AL558929 1040 bp mRNA linear EST 31-MAY-2003
LOCUS AL558929 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DJ007YK05 5-PRIME, mRNA sequence.
ACCESSION AL558929
VERSION AL558929.2 GI:31283062
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1040)
I1,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12903930.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by life technologies, a
division of invitrogen. This sequence belongs to sequence cluster
9238.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ007AF03QPL&cluster=9238.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ007AF03QPL.

```

```

FEATURES
    source          Location/Qualifiers
                    1..1040
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="CS0DJ007YK05"
                    /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
                    /clone_1lib="JURKAT"
                    /note="1st strand cDNA was primed with a NotI-oligo(dT)
                    primer. Five prime end enriched, double-strand cDNA was
                    digested with Not I and cloned into the Not I and EcoR V
                    sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT      188 a      341 c      338 g      162 t      11 others
ORIGIN
Query Match      46.2% Score 924; DB 9; Length 1040;
Best Local Similarity 99.2% Pred. No.1.6e-172;
Matches 958; Conservative 2; Mismatches 2; Indels 4; Gaps 3;

QY      1  CTCGCTCCAAAGTTGTGACGCGGAGCGGCTCGGGGTGTCAGCCGCGTCCGAGAGGCC 60
Db      77  CTCGCTCCAAAGTTGTGACGCGGAGCGGCTCGGGGTGTCAGCCGCGTCCGAGAGGCC 136
QY      61  TCTTGGGGGGGGGGGGG--GCGGCTCGGGGGGCGCCCTGAGCAGAAAAAGAGAA 118
Db      137  TCTTGGGGGGGGGGGGGCGGGGCGGGCTCGGGGGGCGCCCTGAGCAGAAAAAGAGAA 196
QY      119  CGAGGCTCGGTCCAGTGGGACCCAGCTCCTTACTCTGTGTGCCAGCCGCTGGCTGTGG 178
Db      197  CGAGGCTCGGTCCAGTGGGACCCAGCTCCTTACTCTGTGTGCCAGCCGCTGGCTGTGG 256
QY      179  CAGGCATTCGCCAGCTCCCGACTGTGACCACTTCTCAGTGTGCTCTCAGCTCTC 238
Db      257  CAGGCATTCGCCAGCTCCCGACTGTGACCACTTCTCAGTGTGCTCTCAGCTCTC 316
QY      239  AGTTTCCCTCTGGGGGCGATGGCGGGGCGAGGCTCTGTGTTTCTGTGGGCAATTCA 298
Db      317  AGTTTCCCTCTGGGGGCGATGGCGGGGCGAGGCTCTGTGTTTCTGTGGGCAATTCA 375
QY      299  CGGCTGTGATTCGTGAGGAAGTAATTCGCGGGGTGAGACCCCGGTTCTCCGAGCCGGA 358
Db      376  CGGCTGTGATTCGTGAGGAAGTAATTCGCGGGGTGAGACCCCGGTTCTCCGAGCCGGA 435
QY      359  -CCGCCCTCCGCTCAGCAGAGATGCCAAGAGGCGCTGGGCCCGGAGCCAGGTAT 418
Db      436  CCGCCCTCCGCTCAGCAGAGATGCCAAGAGGCGCTGGGCCCGGAGCCAGGTAT 495
QY      419  CACCAATGCGAGACACCCGCCCAAGCCAGGGAGCTGGCTTCGCAAGGGGAGCT 478
Db      496  CACCAATGCGAGACACCCGCCCAAGCCAGGGAGCTGGCTTCGCAAGGGGAGCT 555
QY      479  GGTACACATTCCTGAGAGGCTCGGAGAACAAAGACTGTACCGGTCAACACACACG 538
Db      556  GGTACACATTCCTGAGAGGCTCGGAGAACAAAGACTGTACCGGTCAACACACACG 615
QY      539  TGGACAGAGAGGCTCTGTGAGCTGGGCGCTGCGGAGAGGAGGAGCCCTCCGACAG 598
Db      616  TGGACAGAGAGGCTCTGTGAGCTGGGCGCTGCGGAGAGGAGGAGGAGCCCTCCGACAG 675
QY      599  CCGCAAGCTAGCGCTCATCGCGTGTTCACAGGGAAGATCTGGGCGCAGAGGCTGTCA 658
Db      676  CCGCAAGCTAGCGCTCATCGCGTGTTCACAGGGAAGATCTGGGCGCAGAGGCTGTCA 735
QY      659  GCGAGTGCAGCGCCCGAGAGATGAGGTGTTCGCGGAGGAGTCCGCGCACCCCGG 718
Db      736  GCGAGTGCAGCGCCCGAGAGATGAGGTGTTCGCGGAGGAGTCCGCGCACCCCGG 795
QY      719  CGACTACGTCGTGTGAGCTTGTGGCGGAGCTGATCCTACCGGCTGTCGACCG 778
Db      796  CGACTACGTCGTGTGAGCTTGTGGCGGAGCTGATCCTACCGGCTGTCGACCG 854

```

QY 779 CGACGCGACCTCACAATCGATGAGCGCGTGTCTTCTTGCACTCATGAGCATGTGGA 838
|||||
Db 855 CGAGCGCCACCTCACAATCGATGAGCGCGTGTCTTCTTGCACTCATGAGCATGTGGA 914
|||||
QY 839 GCATTACAGCAAGGACAAAGGCGGCTATCTGCAACCAAGCTGGTGAACCAAGCGAACA 898
|||||
Db 915 GCATTACAGCAAGGACAAAGGCGGCTATCTGCAACCAAGCTGGTGAACCAAGCGAACA 974
|||||
QY 899 CGGAGCAAGTGGCGGAGGAGGAGCTGGCGCGGCGGCGGCTGGTACTGAACTGTGAGCA 958
|||||
Db 975 CGGAGCAAGTGGCGGAGGAGGAGCTGGCGCGGCGGCGGCTGGTACTGAACTGTGAGCA 1034
|||||
QY 959 TTTGAC 964
|||||
Db 1035 TTTGAC 1040
|||||

RESULT 4
AL558805 1032 bp mRNA linear EST 31-MAY-2003
LOCUS AL558805 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens CDNA clone CS0DJ015Y001 5-PRIME, mRNA sequence.
ACCESSION AL558805
VERSION AL558805.2 GI:31280603
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1032)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12903685.
Contact: Genoscope - Centre National de Sequencage
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqef@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9238.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ015AH01P1&cluster=9238.r. Contact :
Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ015AH01P1.

FEATURES
source location/Qualifiers
1..1032
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015Y001"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_11b="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 187 a 338 c 334 g 164 t
ORIGIN

Query Match 45.5% Score 909.6 DB 9: Length 1032:
Best Local Similarity 99.1% Pred. No. 1.1e-169:
Matches 943; Conservative 3; Mismatches 2; Indels 4; Gaps 3:

QY 1 CTCGCTCCAAAGTTGTGACGCGCGGACCGCTCGGGGTGTGACCGCGCTCGCGAGGCC 60
|||||
Db 80 CTCGCTCCAAAGTTGTGACGCGCGGACCGCTCGGGGTGTGACCGCGCTCGCGAGGCC 139
|||||
QY 61 TCTTGGGGGGGGGGGGGG--GCGGCTCGGGGGGGGGGGGGCTGAGCAAGAAACAGAGAACA 118
|||||

Db 140 TCTTGG 199
|||||
QY 119 CCAGGCTGGGTCCAGTGGGACCCAGCTCCCTACCTCCCTGCTGCTGCTGCTGCTGCTGCTG 178
|||||
Db 200 CCAGGCTGGGTCCAGTGGGACCCAGCTCCCTACCTCCCTGCTGCTGCTGCTGCTGCTGCTG 259
|||||
QY 179 CAGGCTATTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 238
|||||
Db 260 CAGGCTATTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 319
|||||
QY 239 AGTTTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 298
|||||
Db 320 AGTTTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 378
|||||
QY 299 CGGCTGTGATTTGCTGAGGAATTTTCCCGGGGTGAGCCCGCTTCTTCCAGGCTTGCA 358
|||||
Db 379 CGGCTGTGATTTGCTGAGGAATTTTCCCGGGGTGAGCCCGCTTCTTCCAGGCTTGCA 438
|||||
QY 359 CCCCCCTCCGCTCAGCGAGGATGCCAAGAGGCGCTGGGCGCGGCGCGGCGCGGCGCGGCGG 418
|||||
Db 439 CCCCCCTCCGCTCAGCGAGGATGCCAAGAGGCGCTGGGCGCGGCGCGGCGCGGCGCGGCGG 498
|||||
QY 419 CACCAATGCGAGCACACCCCGCCCAAGCGAGGAGCTGGCTTCCGCAAGGCGGAGCT 478
|||||
Db 459 CACCAATGCGAGCACACCCCGCCCAAGCGAGGAGCTGGCTTCCGCAAGGCGGAGCT 558
|||||
QY 479 GGTCAACATCTTGGAGGCTGGGAGAAAGAGTGTGATCCCGCTCAAGACCAACAG 538
|||||
Db 559 GGTCAACATCTTGGAGGCTGGGAGAAAGAGTGTGATCCCGCTCAAGACCAACAG 618
|||||
QY 539 TGGACAGAGAGGGGCTGTGAGCTGGGCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 598
|||||
Db 619 TGGACAGAGAGGGGCTGTGAGCTGGGCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 678
|||||
QY 599 CCCCAAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 658
|||||
Db 679 CCCCAAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 738
|||||
QY 659 GGAGGTGAGAGCTCCCGAGGATGGGCTGTCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGG 718
|||||
Db 739 GGAGGTGAGAGCTCCCGAGGATGGGCTGTCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGG 798
|||||

QY 719 CGACTACGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 778
|||||
Db 799 CGACTACGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 858
|||||

QY 779 CGAGCGGCACTCACAATCGATGAGCGCGTGTCTTCTTGCACTCATGAGCATGTGGA 838
|||||
Db 859 CGAGCGGCACTCACAATCGATGAGCGCGTGTCTTCTTGCACTCATGAGCATGTGGA 918
|||||

QY 839 GCATTACAGCAAGGACAAAGGCGGCTATCTGCAACCAAGCTGGTGAACCAAGCGAACA 898
|||||
Db 919 GCATTACAGCAAGGACAAAGGCGGCTATCTGCAACCAAGCTGGTGAACCAAGCGAACA 978
|||||

QY 899 CGGAGCAAGTGGCGGAGGAGGAGCTGGCGCGGCGGCGGCTGGTACTGAACTGTGAGCA 950
|||||
Db 979 CGGAGCAAGTGGCGGAGGAGGAGCTGGCGCGGCGGCGGCTGGTACTGAACTGTGAGCA 1029
|||||

RESULT 5
BG744770 906 bp mRNA linear EST 15-MAY-2001
LOCUS BG744770/0 60272289191 NIH_MGC_106 Homo sapiens CDNA clone IMAGE:4849175 3',
DEFINITION mRNA sequence.
ACCESSION BG744770
VERSION BG744770.1 GI:14055423
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 906)
AUTHORS NIH-MGC http://mhc.ncl.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DRS/NCI

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: LNCM688 row: f column: 24
High quality sequence start: 21
High quality sequence stop: 845.
Location/Qualifiers

FEATURES

source

1. 906
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4849175"
/issue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_106"
/note="Organ: Blood; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(C). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 142 a 306 c 281 g 176 t 1 others

ORIGIN

Query Match

Best Local Similarity 43.4%; Score 867; DB 10; Length 906;

Matches 887; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

1031 GGTGGCCGTGAGAAATATCAAGTGTATGTGACAGCCAGCCCTCTCGAGCAGAGCGC 1090
1091 GGTGGCCGTGAGAAATATCAAGTGTATGTGACAGCCAGCCCTCTCGAGCAGAGCGC 848
1091 CGTCATGACGAAGATGCAACAGAGAACTGTGCTCTCTGGGGCTGATCTCTGACCA 1150
847 CGTCATGACGAAGATGCAACAGAGAACTGTGCTCTCTGGGGCTGATCTCTGACCA 788
1151 GGGGCTGTACATGTGATGAGCAGCGAGCAAGGGCAACCTGTGAACTTTCTGCGGAC 1210
787 GGGGCTGTACATGTGATGAGCAGCGAGCAAGGGCAACCTGTGAACTTTCTGCGGAC 728
1211 CGGGGCTGAGCCCTGCTGAACACCGCTCACTCTGCACTTTCTCTGACGTGGCGCA 1270
727 CGGGGCTGAGCCCTGCTGAACACCGCTCACTCTGCACTTTCTCTGACGTGGCGCA 668
1271 GGGCATGAGTACTTGGAGAGCAAGAAAGCTTGTGACCGGACCTGGCCCGCAACAT 1330
667 GGGCATGAGTACTTGGAGAGCAAGAAAGCTTGTGACCGGACCTGGCCCGCAACAT 608
1331 CTGTGCTCAGAGAGACTGTGAGCAAGTGTGACGACTTTGGCTGGCCCAAGCCGAGCG 1390
607 CTGTGCTCAGAGAGACTGTGAGCAAGTGTGACGACTTTGGCTGGCCCAAGCCGAGCG 548
1391 GAAAGGGCTAGACTCAAGCGGCTGCCCTCAAGTGTGAGCGCCGAGGCTCTCAACA 1450
547 GAAAGGGCTAGACTCAAGCGGCTGCCCTCAAGTGTGAGCGCCGAGGCTCTCAACA 488
1451 CGGGAAGTTTACAGCAAGTGTGAGTGTGAGATTTTGGGCTCTGCTGGAGGCTTT 1510
487 CGGGAAGTTTACAGCAAGTGTGAGTGTGAGATTTTGGGCTCTGCTGGAGGCTTT 428
1511 CTATATGAGAGGGCTCCGATCAAAATGTCTAGTGAAGAGAGTGTGAGGCGCTGGA 1570
427 CTATATGAGAGGGCTCCGATCAAAATGTCTAGTGAAGAGAGTGTGAGGCGCTGGA 368

1571 GAAAGGGCTAGACTGAGAACCCCGAGAGGCTGTCCAGGCCCCGTGCACCTCTCATGAG 1630
367 GAAAGGGCTAGACTGAGAACCCCGAGAGGCTGTCCAGGCCCCGTGCACCTCTCATGAG 308
1631 CAGCTGTGTGGAGGACAGAGCCCGCCGCGCCGACCTTCCGAACATGGCCAGAGCT 1690
307 CAGCTGTGTGGAGGACAGAGCCCGCCGCGCCGACCTTCCGAACATGGCCAGAGCT 248
1691 GGGCCGGAGACTACGAGTGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1750
247 GGGCCGGAGACTACGAGTGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 188
1751 CTCACCTGCTCCCGAGAGCCAGAGCCCTGACCCACCCCGGTGGGCTTGGCCCGAGA 1810
187 CTCACCTGCTCCCGAGAGCCAGAGCCCTGACCCACCCCGGTGGGCTTGGCCCGAGA 128
1811 GGAACGAGAGAGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1870
127 GGAACGAGAGAGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 68
1871 GGGGCAAGTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1930
67 GGGGCAAGTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8
1931 GGGC 1934
7 GGGC 4

RESULT 6

AL580543/c

AL580543 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED

DEFINITION Homo sapiens cDNA clone CSDBD015Y001 3-PRIME, mRNA sequence.

ACCESSION AL580543

VERSION AL580543.2 GI:31318812

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1012)

AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polyes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT On Feb 16, 2001 this sequence version replaced gi:12946667.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 9238.r For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CSDBD015AH01NP1&cluster=9238.r. Contact :

Feng Liang Email: fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CSDBD015AH01NP1.

location/Qualifiers

1. 1012

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSDBD015Y001"

/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"

/clone_lib="JURKAT"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 158 a 346 c 277 g 194 t 37 others
ORIGIN

Query Match 42.9% Score 857.2: DB 9: Length 1012;
Best Local Similarity 92.6% Pred. No. 2.6e-159;
Matches 915: Conservative 14: Mismatches 55: Indels 4: Gaps 4:

```

OY 984 GAGGAGAGATTGGAGCTGTCTGTGAGGAGTACTGTGGGCAAAAGAGCCGTGAAG 1043
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 986 GAGAGAGAGATTGGAGCTGTCTGTGAGGAGTACTGTGGGCAAAAGAGCCGTGAAG 929
OY 1044 AATATCAAGTGTATGTGACAGCCAGGCGCTTCTGTGACGAGACGCGGTATGACGAAG 1103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 928 AATATCAAGTGTATGTGACAGCCAGGCGCTTCTGTGACGAGACGCGGTATGACGAAG 869
OY 1104 ATGCAACAGAGAACTGTGCGCTCTGTGGGCGGTGATCCCTGACACAGGCGGTGACATT 1163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 868 ATGCAACAGAGAACTGTGCGCTCTGTGGGCGGTGATCCCTGACACAGGCGGTGACATT 809
OY 1164 GTCATGAGCAGCTGTGACAGAGGCGCAACCTGTGTAACCTTCTGCGGACCGGGGTGACACC 1223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 808 GTCATGAGCAGCTGTGACAGAGGCGCAACCTGTGTAACCTTCTGCGGACCGGGGTGACACC 749
OY 1224 CTCTGGAACACCGCTCAGCTCTCTGCAAGTTTCTCTGACAGTGGCCGAGGCGATGAGTAC 1283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 748 CTCTGGAACACCGCTCAGCTCTCTGCAAGTTTCTCTGACAGTGGCCGAGGCGATGAGTAC 689
OY 1284 CTGGAGACGAAGAAGCTGTGACACCGGCGCAACCTGTGGGCGCGCCGCAACATCTGCTCAGAG 1343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 688 CTGGAGACGAAGAAGCTGTGACACCGGCGCAACCTGTGGGCGCGCCGCAACATCTGCTCAGAG 629
OY 1344 GACCTGTGTGGCCAAAGTGTGAGGACTTTGGCTGTGGCCCAAGGCGAGCGAGGCGCTAAGC 1403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 628 GACCTGTGTGGCCAAAGTGTGAGGACTTTGGCTGTGGCCCAAGGCGAGCGAGGCGCTAAGC 569
OY 1404 TCAAGCCGGCTGCCCGTCAAGTGTGAGGCGCGCCGCAAGGCTCTCAACACAGGGAAGTTACCC 1463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 568 TCAAGCCGGCTGCCCGTCAAGTGTGAGGCGCGCCGCAAGGCTCTCAACACAGGGAAGTTACCC 509
OY 1464 AGCAAGTGTGAGTGTGAGTGTGGGGTGTGCTGGGAGGCTTCTCATATATGAGAGG 1523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 508 AGCAAGTGTGAGTGTGAGTGTGGGGTGTGCTGGGAGGCTTCTCATATATGAGAGG 449
OY 1524 GCTCCGTACCTTAATAATGTCACTGAAAGAGGTGTGAGGCGCGGTGAGAGAGGGGTACCGC 1583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 448 GCTCCGTACCTTAATAATGTCACTGAAAGAGGTGTGAGGCGCGGTGAGAGAGGGGTACCGC 389
OY 1584 ATGGAACCCCGGAGGCGTGTCCAGGCGCCCGTGCACGTCCTCATAGAGCAGCTGTGGAG 1643
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 388 ATGGAACCCCGGAGGCGTGTCCAGGCGCCCGTGCACGTCCTCATAGAGCAGCTGTGGAG 329
OY 1644 GCAGAGCCCGGCGGCGGCGCAACCTTCCGCAAAAGCGGCGAAGAGTGGCCCGGAGCTA 1703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 328 GCAGAGCCCGGCGGCGGCGCAACCTTCCGCAAAAGCGGCGAAGAGTGGCCCGGAG-7A 270
OY 1704 GCAGTGTGAGTGTGCCACGCTCTGTCTCAGGCGCAGG-ACGCGCAGCGGCTCTCACCTCCGC 1762
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 269 CGCAGTGTGAGTGTGCCACGCTCTGTCTCAGGCGCAGGAGCAGCCGACRCTCTCACCTCCGCY 210
OY 1763 CCGAAGCCAGAGAGCCCTGACCCACCGGCTGGGGCCCTTGGCCCGCCAGAGAGCAGAGAG 1822
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 209 TCGAAKCCAGAGAGCTGTACMCACCCGCTGGGGCTCTTGGCCCGCCAGAGAGCAGAGAG 150
OY 1823 TGGAGAGTGTGGGGGCGTGTGACAGGCGCCCAAGGAGGAGGTCCAGGCGGCGCAAGTAT 1882
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 149 TGGAGAGTGTGGGGGCGTGTGACAGGCGCCCAAGGAGGAGGAGGCGCGCAAGTAT 90
OY 1883 CCTCTGTGTGCCACAGAGAGGCGCTGAGCCACGTAGGGGCGCTGTGGGCGCGCGGTGACA 1942
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 89 CTTCTGTGGCGCCACAGAGAGGAGGCGGCGCACAGTATAGGGGSKNNNGGGGSGGNNNN 30
OY 1943 CCCCAAGACCTCGGAAGATGTGCGCCG 1970
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

DB 29 GNNNGGGGNGTGTGNGGTGNNKNNNTG 2

RESULT 7
BO054024 1043 bp mRNA linear EST 29-MAR-2002
AGENCOURT 6874039 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5935995
DEFINITION 5', mRNA sequence.
ACCESSION BO054024
VERSION BO054024.1 GI:19813364
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 1043)
NIH-MGC <http://mhc.ncl.nih.gov/>.
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI

cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: L10M2124 row: k column: 04
High quality sequence stop: 616.

FEATURES
source
1..1043
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5935995"
/tissue="type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_106"
/note="Organ: blood; Vector: pOTB1; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

BASE COUNT 225 a 310 c 337 g 170 t 1 others
ORIGIN

Query Match 42.3% Score 845.4: DB 12: Length 1043;
Best Local Similarity 98.8% Pred. No. 5.6e-157;
Matches 883: Conservative 0: Mismatches 7: Indels 4: Gaps 3:

```

OY 379 GGATGCCAAGAGCGGTGGGCGCCCGGCGCACCAAGTGTATCAACCAATGAGAGACACCC 438
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 GGAATGCCAAGAGCGGTGGGCGCCCGGCGCACCAAGTGTATCAACCAATGAGAGACACCC 60
OY 439 GCCCCAGCAGAGGAGCTGCGCTTCCGCAAGGCGAGCGTGTACCATCTGTGAGGCT 498
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 GCCCCAGCAGAGGAGCTGCGCTTCCGCAAGGCGAGCGTGTACCATCTGTGAGGCT 120
OY 499 GCGAGAACAAAGAGCTGTACCGCGCTCAAGCAGCAGCAGTGTGACAGTGTGAGGCT 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 121 GCGAGAACAAAGAGCTGTACCGCGCTCAAGCAGCAGCAGTGTGACAGTGTGAGGCT 180
OY 559 CAGCTGGGCGCTCGGAGGCGGAGGCGCTCTCCGAGAGCAGCCCAAGCTCAAGCTCATAGC 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 181 CAGCTGGGCGCTCGGAGGCGGAGGAGCGCTCTCCGAGAGCAGCCCAAGCTCAAGCTCATAGC 240
OY 619 CGTGGTTCCACGGGAAGATCTCGGGCCAGAGAGCTGTTCACAGCAGCTGACGCTCCGAGG 678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 241 CGTGGTTCCACGGGAAGATCTCGGGCCAGAGAGCTGTTCACAGCAGCTGACGCTCCGAGG 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

OY 679 ATAGGCTGTTCTGCTGCTGCGGAGTCCGGCGCCACCCCGGAGTACGTCTGTGCTGA 728
    |||
DB 301 ATGGGCTGTTCTGCTGCTGCGGAGTCCGGCGCCACCCCGGAGTACGTCTGTGCTGA 360
OY 739 GCTTTGGCGCGAGCTATCCACTACCGGCTGTCGACCGCGGACCAGCCACTACATTCG 798
    |||
DB 361 GCTTTGGCGCGAGCTATCCACTACCGGCTGTCGACCGCGGACCAGCCACTACATTCG 420
OY 799 ATGAGGCGCTGTTCTTCTGCAACCTCATGAGCATGTGAGCATTTACAGCAAGAGCAAG 858
    |||
DB 421 ATGAGGCGCTGTTCTTCTGCAACCTCATGAGCATGTGAGCATTTACAGCAAGAGCAAG 480
OY 859 GCGCTATCTGCACCAAGCTGTGAGACCAAAAGGGAACACGGGACCAAGTCCGGCAGG 918
    |||
DB 481 GCGCTATCTGCACCAAGCTGTGAGACCAAAAGGGAACACGGGACCAAGTCCGGCAGG 540
OY 919 AGGAGCTGCGCAGGCGCGGCTGTTACTGAACTGACGATTTGACATTTGGAGCAAGCA 978
    |||
DB 541 AGGAGCTGCGCAGGCGCGGCTGTTACTGAACTGACGATTTGACATTTGGAGCAAGCA 600
OY 979 TCGGAGAGGAGAGTTTGAAGCTGTCTGAGGGTGAATCTGCGGCAAAAGTGGCCG 1038
    |||
DB 601 TCGGAGAGGAGAGTTTGAAGCTGTCTGAGGGTGAATCTGCGGCAAAAGTGGCCG 660
OY 1039 TGAAGATATCAAGTGTGATGTGACAGACCCAGGCTTCTGAGACGAGACCGGCTCATGA 1098
    |||
DB 661 TGAAGATATCAAGTGTGATGTGACAGACCCAGGCTTCTGAGACGAGACCGGCTCATGA 720
OY 1099 CGAAGATGCAACAGAG- AACCTGTCGCTCTCTGCGGCTGATTCGCAACAGGGGCTG 1157
    |||
DB 721 CGAAGATGCAACAGAGAACCTGTCGCTCTCTGCGGCTGATTCGCAACAGGGGCTG 780
OY 1158 TACATTTGTCATGAGACAGTGAACAGGCAACCT- GGTGAACCTTTCTGCGGACCGGCGG 1216
    |||
DB 781 TACATTTGTCATGAGACAGTGAACAGGCAACCTGGTGAACCTTTCTGCGGACCGGCGG 840
OY 1217 --TCGAGCCCTGTCGAACACCGCTCAGCTCTCTGCAAGTTTCTGCAAGTGGCC 1268
    |||
DB 841 GTCGAGCCCTGTCGAACACCGCTCAGCTCTCTGCAAGTTTCTGCAAGTGGCC 894

RESULT 8
AL567349/c 949 bp mRNA linear EST 31-MAY-2003
LOCUS AL567349 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS0DF027YD13 3-PRIME, mRNA sequence.
ACCESSION AL567349
VERSION AL567349.2 GI:31290254
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homini; Homo.
1 (bases 1 to 949)
L1,W.B., Gruber,C., Jessee,J. and Polyes,D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 16, 2001 this sequence version replaced gi:12920618.
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: sequef@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9238.r for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgl-bin/cluster.cgi?seq=CS0DF027CB07NP1&cluster=9238.r. Contact :
Feng Liang, Email: fliang@life.techn.com URL :
http://fulllength.invitrogen.com/
Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DF027CB07NP1.
FEATURES
source 1..949

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF027YD13"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/Note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
BASE COUNT 144 a 279 c 290 g 212 t 24 others
ORIGIN
Query Match 41.5%; Score 830; DB 9; Length 949;
Best Local Similarity 93.3%; Pred. No. 6e-154;
Matches 876; Conservative 18; Mismatches 42; Indels 3; Gaps 3;

OY 1013 TGAATACCTGGGCAAAAGGTGGCCGTGAAGAAATCAAGTGTGATGACAGCCAGGC 1072
    |||
DB 949 TGAATACCTGGGCAAAAGGTGGCCGTGAAGAAATCAAGTGTGATGACAGCCAGGC 890
OY 1073 CTTCCGAGACAGACGGCCCTCATGACGAGATGCAACACGAGAACCTGGTGGCTCCT 1132
    |||
DB 889 CTTCTGAGAGAGAGCGGCTCATGACGAGATGCAACACGAGAACCTGGTGGCTCCT 830
OY 1133 GGGCGTATCCTCAGCAGGCGCTGATCATTTGATGAGACAGTGAAGGCAACCT 1192
    |||
DB 829 GGGCGTATCCTCAGCAGGCGCTGATCATTTGATGAGACAGTGAAGGCAACCT 770
OY 1193 GGTGAACCTTTCTGCGGACCGGCGGTGAGCCCTGTCGAACACCGCTCAGCTCTGAGTT 1252
    |||
DB 769 GGTGAACCTTTCTGCGGACCGGCGGTGAGCCCTGTCGAACACCGCTCAGCTCTGAGTT 710
OY 1253 TTCTCGACAGTGGCGGAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 1312
    |||
DB 709 TTCTCGACAGTGGCGGAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 650
OY 1313 CTTGCGGCGCGGCAACATCTGCTTCAGAGACCTGTGAGGCAACGAGTGGAGCTTTGG 1372
    |||
DB 649 CTTGCGGCGCGGCAACATCTGCTTCAGAGACCTGTGAGGCAACGAGTGGAGCTTTGG 590
OY 1373 CTTGCGGCGCGGCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1432
    |||
DB 589 CTTGCGGCGCGGCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 530
OY 1433 GCGCGAGGCTCTCAACACGCGGAGGAGTTCACCAAGTGGAGTGGAGTGGAGTGGAGTGG 1492
    |||
DB 529 GCGCGAGGCTCTCAACACGCGGAGGAGTTCACCAAGTGGAGTGGAGTGGAGTGGAGTGG 470
OY 1493 GCTGCTTGGGAGGCTTCTCATATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1552
    |||
DB 469 GCTGCTTGGGAGGATCTTCTCATATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 410
OY 1553 GGTGTGAGAGCGGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1612
    |||
DB 409 GGTGTGAGAGCGGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 350
OY 1613 GGTGTGAGAGCGGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1672
    |||
DB 349 GGTGTGAGAGCGGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 290
OY 1673 CAACACTGGCGGAGAACTGGCCGGGAGGAGTACGACGAGTGGAGGAGGAGGAGGAGGAGGAG 1732
    |||
DB 289 CAACACTGGCGGAGAACTGGCCGGGAGGAGTACGACGAGTGGAGGAGGAGGAGGAGGAGGAG 231
OY 1733 AGGCGAGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1792
    |||
DB 230 AGGCGAGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 171
OY 1793 GGGGCGGCTTGGCGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1852
    |||

```

Db 170 -GGGSCCTTGGCCCCAGAGACGAGAGAGTGAGAGTGCAGCGTGGGGGACATGTACCAG 112

Qy 1853 GCCAAGAGAGGATGCAGCGGGGCAAGTATCCTCTCGTGGTCCACAGAGGGGCGTGGCC 1912

Db 111 GCCCAAGAGAGGCCCCAGCGAGCAAGTACACCMCT-CTGCCACAGGAGGCTGNCCTCA 53

Qy 1913 ACGTAGGGGCTCTGGCGGCGCGTGGACACCCAGAC 1951

Db 52 CGTNAGGGGNCCTGTGGCGGCGCGTGGACACCCAGAC 14

RESULT 9
BMS61134 1102 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT.6566231 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5550935
DEFINITION 5', mRNA sequence.
ACCESSION BMS61134
VERSION BMS61134.1 GI:18806135
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1102)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strassberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLM12263 row: n column: 24
High quality sequence start: 32
High quality sequence stop: 611.
Location/Qualifiers
1. 1102
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5550935"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo
Average insert size 1.75 kb. Library constructed by Life
Technologies."

BASE COUNT 199 a 377 c 354 g 172 t

ORIGIN

Query Match 40.6%; Score 811.4; DB 12; Length 1102;
Best Local Similarity 95.6%; Pred. No. 3e-150;
Matches 856; Conservative 0; Mismatches 36; Indels 3; Gaps 2;

Qy 18 AGCCGGAGCCGCTCGGGGTGTGACGCCGCTCGGAGAGCCCTCTGGGGGCGGGCGG 77

Db 44 AGCCGGAGCCGCTCGGGGTGTGACGCCGCTCGGAGAGCCCTCTGGGGGCGGGCGG 103

Qy 78 GG--GGCGCTGGGGGGCGGGCGCGCTGAGAGAAAGAGAAAGAACACAGGCTGGTCAAGT 135

Db 104 GGGCGGGCTCGGGGGCGGGCGCGCTGAGAGAAAGAGAAAGAACACAGGCTGGTCAAGT 163

Qy 136 GCACCAAGCTCCCTACTCTGTGTGACAGCCGCTGGCTGTGACAGGCAATTCACAGGT 195

Db 164 GCACCAAGCTCCCTACTCTGTGTGACAGCCGCTGGCTGTGACAGGCAATTCACAGGT 223

Qy 196 CCGGCACTGTACCACTTGTCAAGTGTGCTCTACCTGCTCAAGTTTCCTCTGGGGGG 255

Db 224 CCCCAGTGTGACACACTTGTCTAGTGTGCTCTCTACCTGCTCAAGTTTCCCTCT-GGGG 282

Qy 256 CGATGGGGGGGCGAGGCTCTCTGTTTCTTGGCGGGGCAATTCACGGCTGTATTTGTGTG 315

Db 283 CGATGGCGGGTCTAGGCTCTCTGTTTCTTGGCGGGGCAATTCACGGCTGTATTTGTGTG 342

Qy 316 AGAATCTTCCCCGGGTGAGCCCCCGCTTCTCTCCAGAGCTGGACACCCCTCCGCTAG 375

Db 343 AGAATCTTCCCCGGGTGAGCCCCCGCTTCTCTCCAGAGCTGGACACCCCTCCGCTAG 402

Qy 376 CCAGATGCCAAGCAGAGCGCTGGGCGCCCGGGACCAAGTATTCACCAATTCAGACACA 435

Db 403 CCAGATGCCAAGCAGAGCGCTGGGCGCCCGGGACCAAGTATTCACCAATTCAGACACA 462

Qy 436 CCGCGCCCAAGCCAGAGGAGCTGGCTTCCGCAAGGGGAGAGTGTATCATCTCTGAGG 495

Db 463 CCGCGCCCAAGCCAGAGGAGCTGGCTTCCGCAAGGGGAGAGTGTATCATCTCTGAGG 522

Qy 496 CCGCGCAAGACAGAGCTGTACCGCTCAAGCAGCAGCAGTGGAGAGAGGGGCTGC 555

Db 523 CCGCGCAAGACAGAGCTGTACCGCTCAAGCAGCAGCAGTGGAGAGAGGGGCTGC 582

Qy 556 TGGCAGCTGGGGCGCTGGGGAGCGGAGGCCCTTCTCGCAGACCCCAAGCTCAGCTCA 615

Db 583 TGGCAGCTGGGGCGCTGGGGAGCGGAGGCCCTTCTCGCAGACCCCAAGCTCAGCTCA 642

Qy 616 TGCCGTGGTTCCAGCGGGAAGATCTGGGGCCAGAGAGCTGTCCAGAGCTCAGCTCCG 675

Db 643 TGCCGTGGTTCCAGCGGGAAGATCTGGGGCCAGAGAGCTGTCCAGAGCTCAGCTCCG 702

Qy 676 AGATGGGCTTTCCTGTGGGGAGAGTCCGGCGCCAGCCCGGCACTACGCTGTGTCG 735

Db 703 AGATGGGCTTTCCTGTGGGGAGAGTCCGGCGCCAGCCCGGCACTACGCTGTGTCG 762

Qy 736 TGAGCTTGGCGCGAGCTCATTCACATCCGCTGTCCAGCGGAGCGGCACTCAGCA 795

Db 763 TGAGCTTGGCGCGAGCTCATTCACATCCGCTGTCCAGCGGAGCGGCACTCAGCA 822

Qy 796 TCGATGAGCGCGCTGTCTTCTGCAACCTCATGAGCATGTGAGCATTCAGCAAGGACA 855

Db 823 TCGAAGAGCGCGTGTCTTCTGCAACCTCATGAGCATGTGAGCATTCAGCAAGGAAA 882

Qy 856 AGGCGCTATCTGCACCAAGCTGTGAGACCAACGGAACACGAGCAAGTC 910

Db 883 GCGGCTATCTCTGCCCCAGCTTGGGAGACAAAGTGGGAACAGCAAGAC 937

RESULT 10
AL538511 967 bp mRNA linear EST 31-MAY-2003
LOCUS AL538511 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS0DF027YD13 5-PRIME, mRNA sequence.
ACCESSION AL538511
VERSION AL538511.2 GI:31263101
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 967)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayars, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 13, 2001 this sequence version replaced gi:12802004.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9238.r For
more information about this cluster, see
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DF027YD13>
cgl-bin/cluster.cgi?seq=CS0DF027YD13

Feng Liang Email: fliang@lifetech.com URL: <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID: CS0DF027CB070P1.
Location/Qualifiers
1. .967

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF027YD13"
/tissue_type="FETAL BRAIN"
/dev_stage="Fetal"
/note="Organ: brain; Vector: pCMVSPORT-6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

BASE COUNT 193 a 297 c 330 g 147 t
ORIGIN

Query Match 40.0%; Score 799; DB 9; Length 967;
Best Local Similarity 97.0%; Pred. No. 8.1e-148;
Matches 814; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
287 GCGGGCATTTACGGGCTGTGATTTCTGCTGAGGAATTCCTCCCGGTGAGCCCGCTTCCT 346
129 GGGGGCGGG 188
347 CCGAGCTGGACACCCCTCCCTCCCTCTCTGAGCAGAGATCCAAAGAGGGCTGGGGCCGGG 406
189 CCGAGCTGGACACCCCTCCCTCCCTCTCTGAGCAGAGATCCAAAGAGGGCTGGGGCCGGG 248
407 CACCAATGTATACCAATGCGACACACCCCGCCCAAGCCAGGAGAGCTGGCTTCGG 466
249 CACCAATGTATACCAATGCGACACACCCCGCCCAAGCCAGGAGAGCTGGCTTCGG 308
467 CAAGGGGAGCTGTGTCATCTCGAGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 526
309 CAAGGGGAGCTGTGTCATCTCGAGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 368
527 GCACACACAGTGGAG 586
369 GCACACACAGTGGAG 428
587 CCTCTCCGAG 646
429 CCTCTCCGAG 488
647 GAGAGCTGTGACAGAGCTGAGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 706
489 GAGAGCTGTGACAGAGCTGAGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 548
707 GCGCCACCCCGGGGAGACTAGCTGCTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 766
549 GCGCCACCCCGGGGAGACTAGCTGCTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 608
767 CGTGTGTGACAG 826
609 CGTGTGTGACAG 668
827 GAGAGAGTGGAGAGATTAACAG 886
669 GAGAGAGTGGAGAGATTAACAG 728
887 AAAGCGGAAACAG 946
729 AAAGCGGAAACAG 788
947 GAACTGTGACAGATTTGACATTTGGAGACAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1006
789 GAACCTGTGACAGATTTGACATTTGGAGACAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAG 848
1007 GCAGGGTGAATACCTGGGGCAAAAGGTGGCCGTGAAGAGATATCAAGTGTGATGACAGC 1066

Db 849 GCAGGGTGAATACCTGGGGCAAAAGGTGGCCGTGAAGAGATATCAAGTGTGATGACAGC 908
Oy 1067 CCAGGCTTCTCTGAGCAG 1125
Db 909 CCAGGCTTCTCTGAGCAG 967

RESULT 11
BI768288
LOCUS
DEFINITION
603056458P1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5205831 5',
BI768288
VERSION
BI768288.1 GI:15759866
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 862)
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L16M1516 row: c column: 16
High quality sequence stop: 780.

FEATURES
source
Location/Qualifiers
1. .862

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5205831"
/lab_host="DH108"
/clone_lib="NIH_MGC_122"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24-week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH-MGC Library."

BASE COUNT 143 a 298 c 281 g 140 t
ORIGIN

Query Match 39.8%; Score 795.4; DB 12; Length 862;
Best Local Similarity 98.1%; Pred. No. 4e-147;
Matches 847; Conservative 0; Mismatches 11; Indels 5; Gaps 4;
Oy 30 CTGCGGCTGTGACAGCGGCTGCGGAGAGCCCTCTGGGGGGCGCGGG--CGGGCTCG 87
Db 1 CTGCGGCTGTGACAGCGGCTGCGGAGAGCCCTCTGGGGGGCGCGGGCGGCTCG 60
Oy 88 GGGGGCGCCCTGTGAGCAAGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 147
Db 61 GGGGGCGCCCTGTGAGCAAGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Oy 148 CTACCTCTGTGTGACAGCGGCTGCTGTGAGAGCATTCAGAGCTCCCGAGCTGTGA 207
Db 121 CTACCTCTGTGTGACAGCGGCTGCTGTGAGAGCATTCAGAGCTCCCGAGCTGTGA 180
Oy 208 CCACTTGTCTAGTGTCTCTCTCAGCTCAGCTTCCCTGTGGGGGGGAGATGGCGGGC 267

```

|||||
Db 181 CCACTTGCTAGTGTGCTCTCACTGCTCATGTTCCCTCT-6GGGGCGATGCGGGGC 239
QY 268 GAGGCTCTGTTCTTCTGCGGGCATTTACAGGCTGTGATTCGTGAGGAATTTCCC 327
Db 240 GAGGCTCTGTTCTTCTGCGGGCATTTACAGGCTGTGATTCGTGAGGAATTTCCC 299
QY 328 GGGGAGACCCCGCTTCTCTCGAGAGCTGGACACCCCTCCGCTGTCAGCCAGATGCCAA 387
Db 300 GGGGAGACCCCGCTTCTCTCGAGAGCTGGACACCCCTCCGCTGTCAGCCAGATGCCAA 359
QY 388 CGAGGCTGTGGGCGCGCGGACCCAGTGTATCACCAATGAGAGACACCCGCGCCCAAGC 447
Db 360 CGAGGCTGTGGGCGCGCGGACCCAGTGTATCACCAATGAGAGACACCCGCGCCCAAGC 419
QY 448 CAGGGAGCTGGCTTCCGCAAGGGGCACTGGTCACTATCTTGAGGCTTCCGAGACA 507
Db 420 CAGGGAGCTGGCTTCCGCAAGGGGCACTGGTCACTATCTTGAGGCTTCCGAGACA 479
QY 508 AGAGCTGTACCGGCTCAGACACACACAGTGGACAGAGGGGCTGTGGACAGTGGGG 567
Db 480 AGAGCTGTACCGGCTCAGACACACACAGTGGACAGAGGGGCTGTGGACAGTGGGG 539
QY 568 CGCTGCGGAGCGGGAGCGCCCTCTCCGAGACCCCAAGCTCAGCTCATGCGCTGTCC 627
Db 540 CGCTGCGGAGCGGGAGCGCCCTCTCCGAGACCCCAAGCTCAGCTCATGCGCTGTCC 599
QY 628 ACAGGAAGATCTCGGGGCGAGAGGCTGTCCAGAGCTGCAAGCTTCCGAGAGTGGCTGT 687
Db 600 ACAGGAAGATCTCGGGGCGAGAGGCTGTCCAGAGCTGCAAGCTTCCGAGAGTGGCTGT 659
QY 688 TCCGTGTGCGGAGTCTCGCGGCGCACCCCGGCGACTGCTGCTGTGGCTGTGGCTTGGCC 747
Db 660 TCCGTGTGCGGAGTCTCGCGGCGCACCCCGGCGACTGCTGCTGTGGCTGTGGCTTGGCC 719
QY 748 GCGAGCTATCTCACTACCGGCTGTGCAACGCGGACGCGCACTGCAATGATGAGGCGG 807
Db 720 GCGAGCTATCTCACTACCGGCTGTGCAACGCGGACGCGCACTGCAATGATGAGGCGG 779
QY 808 TGTCTCTCTGCAACCTATGGA-CATGCTGAGAGCATTAACAGCA-GGACAGAGGCGCTAT 865
Db 780 TGTCTCTCTGCAACCTATGGA-CATGCTGAGAGCATTAACAGCA-GGACAGAGGCGCTAT 839
QY 866 CTCGACCAAGCTGTGAGACCAA 888
Db 840 CTCGACCAAGCTGTGAGACCAA 862

RESULT 12
BI819535 912 bp mRNA linear EST 04-OCT-2001
LOCUS 603036746F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5177986 5',
DEFINITION mRNA sequence.
ACCESSION BI819535
VERSION BI819535.1 GI:15931085
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: rs9aps-femail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov

```

```

FEATURES
source
location/Qualifiers
1..912
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5177986"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb.
Insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH MGC Library."

BASE COUNT 153 a 310 c 299 g 150 t
ORIGIN
Query Match 38.1%; Score 762; DB 12; Length 912;
Best Local Similarity 96.1%; Pred. No. 1.6e-140;
Matches 867; Conservative 0; Mismatches 25; Indels 10; Gaps 8;

QY 1 CTCGCTCCAAAGTTGTGCAAGCCGGGAGCCCTCGGGGTGTGACGCGGCTGCGGAGGCC 60
Db 15 CTCCTCTCAAGTTGTGCAAGCCGGGAGCCCTCGGGGTGTGACGCGGCTGCGGAGGCC 74
QY 61 TCCGTGGGGGGGGGGGGGG--GGGGCTGGGGGGGGCCCTGAGCAGAAAAAGAGAGAA 118
Db 75 TCCGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 134
QY 119 CAGAGCTGCGTCAAGTGGACACCAAGCTCCCTGCTGTGTCAGCGGCTGTGGCTGTGG 178
Db 135 CAGAGCTGCGTCAAGTGGACACCAAGCTCCCTGCTGTGTCAGCGGCTGTGGCTGTGG 194
QY 179 CAGGCAATTTGCC-AGCGTCCCGGACAGTGTGACCACTGTGCTAGTGTGCTCTACCTGCT 237
Db 195 CAGGCAATTTGCCCTGTGCGCTGTGCGGCTGTGACCACTGTGCTAGTGTGCTCTACCTGCT 254
QY 238 CAGTTCCCTGCTGGGGGGGGGAGTGGGGGGGGAGGCTGTGCTGTGCTGTGGGGGCAATTTTC 297
Db 255 CAGTTCCCTGCT-6GGGGCGATGGGGGGGGAGGCTGTGCTGTGCTGTGGGGGCAATTTTC 313
QY 298 ACGGCTGTGATTTCTGTGAGAACTTCCCGGGGTGAGCCGCCGCTTCTCGAGACTGAGC 357
Db 314 ACGGCTGTGATTTCTGTGAGAACTTCCCGGGGTGAGCCGCCGCTTCTCGAGACTGAGC 373
QY 358 ACCCCCTCCCGTCTACAGCAGATGCCAAGAGGCGCTGGGCCCCGGGACCCAGGTGTA 417
Db 374 ACCCCCTCCCGTCTACAGCAGATGCCAAGAGGCGCTGGGCCCCGGGACCCAGGTGTA 433
QY 418 TCACCAATGGAGACACACCGCGCCCAAGCGAGGGAGGCTGTGCTGTGCGAAGGGCGAGG 477
Db 434 TCACCAATGGAGACACACCGCGCCCAAGCGAGGGAGGCTGTGCTGTGCGAAGGGCGAGG 492
QY 478 TGTGTACCAATCTTGAGAGCGCTGCGAGAAAGAGCTGTACCGGCTCAAGCAGCAGACA 537
Db 493 TGTGTACCAATCTTGAGAGCGCTGCGAGAAAGAGCTGTACCGGCTCAAGCAGCAGACA 552
QY 538 GTGACACAGAGAGGGCTGTGCGAGCTGGGGCGCTGCGGAGAGCGGGAGGCGCTTCTCCGAG 597
Db 553 GTGACACAGAGAGGGCTGTGCGAGCTGGGGCGCTGCGGAGAGCGGGAGGCGCTTCTCCGAG 612
QY 598 ACCCAAGCTCAGGCTCATGCGCGGTTCACAGGAGAGATCTGGGCGAGAGGCTGTGCG 657
Db 613 ACCCAAGCTCAGGCTCATGCGCGGTTCACAGGAGAGATCTGGGCGAGAGGCTGTGCG 672
QY 658 AGCAGCTGACGCTCCCGAGAGATGGCTGTCTGTGCTGTGCGGAGATCTGCGGCGCACCCCG 717

```

Db 673 AGCAGCTGCAGCTCCCGAGATGAGGCTGTTCTCTGTCGGAGTCCG-GCCGACCCCG 731
Qy 718 GCGACTACGTCCTGTCGCTGAGCTTTGGCCGCGAGC-TCATCCACTACCGCGTGTGCAC 776
Db 732 GCGACTACGTCCTGTCGCTGAGCTTTGGCCGCGAGTTCATCCACTACCGCGTGTGCAC 791
Qy 777 CCGGACGGCCACTCACAATGATGAGCCGCTGTTCTTTCGCAACCTCATGACATGCTG 836
Db 792 --CGGACGGCCACTCACAATGATGAGCCGCTGTTCTTTCGCAACCTCATGACATGCTG 849
Qy 837 GAGCATTTACA-GCAAGACAAAGGCGCTATTCGACCAAGCTGTGAGACCAAGAGGGA 895
Db 850 GAGCATTTACAGCAGGAGGAGGCGGCGCATTCGACCAAGCTGTGAGACCAAGAGGGA 909
Qy 896 AC 897
Db 910 AC 911
RESULT 13
LOCUS BM925586 1014 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT.6710058 NIH_MGC_114 Homo sapiens CDNA clone IMAGE:5763792
5' mRNA sequence.
ACCESSION BM925586
VERSION BM925586.1 GI:19375965
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1014)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12816 row: d column: 01
High quality sequence start: 7
High quality sequence stop: 713.
Location/Qualifiers
1..1014
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5763792"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains; age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH-MGC library."
BASE COUNT 164 a 355 c 330 g 162 t 3 others
ORIGIN
Query Match 36.9%; Score 737.2; DB 12; Length 1014;
Best Local Similarity 96.0%; Pred. No. 1.4e-135;
Matches 842; Conservative 0; Mismatches 24; Indels 11; Gaps 8;
Qy 1 CTCGGTCCAGTTGTGCGACCGCGGACCGCTCTGGGGGTGTGCGACCGCGCTCGGAGGCC 60
|||||

Db 128 CTCGGTCCAGTTGTGCGACCGCGGACCGCTCTGGGGGTGTGCGACCGCGCTCGGAGGCC 187
Qy 61 TCTTGGGGGCGGGGCGCGGg--GCGGCTGCGGGGGCCGCCCTTACAGAAACAGGAAGA 118
Db 188 TCTTGGGGGCGGGGCGGGGCGGGGCTCGGGGGCCGCCCTTACAGAAACAGGAAGA 247
Qy 119 CCAGGCTGGTCCAGTGGACCCAGCTCCCTACCTCTCTGTGCGACAGCGGCTGGCTGTG 178
Db 248 CCAGGCTGGTCCAGTGGACCCAGCTCCCTACCTCTCTGTGCGACAGCGGCTGGCTGTG 307
Qy 179 GAGGCAATTCGAGGCTCCGAGCTGTGACCACTTGTAGTGTGCTCTACCTGCGCTC 238
Db 308 GAGGCAATTCGAGGCTCCGAGCTGTGACCACTTGTAGTGTGCTCTACCTGCGCTC 367
Qy 239 AGTTTCCTCTGCGGGGCGATGGCGGGGCGAGGCTCTGTGTTCTCTGCGGGCATTTCA 298
Db 368 AGTTTCCTCTCT--GGGGGCGATGGCGGGGCGAGGCTCTGTGTTCTCTGCGGGCATTTCA 426
Qy 299 CGGCTGTGATTTCTGTGAGGAACCTTCCCGGGTGAAGCCCGCTTCTCCGAGCCTGGCA 358
Db 427 CGGCTGTGATTTCTGTGAGGAACCTTCCCGGGTGAAGCCCGCTTCTCCGAGCCTGGCA 486
Qy 359 CCGCCCTCCGCTCAGCCAGGATGCCAAGAGAGGCGCTGGGCGCGGGGACCCAGGTAT 418
Db 487 CCGCCCTCCGCTCAGCCAGGATGCCAAGAGAGGCGCTGGGCGCGGGGACCCAGGTAT 546
Qy 419 CACCAATTCGAGCAGCACACCCGCGCAAGGCGAGGAGCTGCGCTTCCGCAAGGCGAGCT 478
Db 547 CACCAATTCGAGCAGCACACCCGCGCAAGGCGAGGAGCTGCGCTTCCGCAAGGCGAGCT 606
Qy 479 GGTACACATCTGTGAGGCTGTGAGAACAGAGCTGTGACCGGCTCAGACACACACAG 538
Db 607 GGTACACATCTGTGAGGCTGTGAGAACAGAGCTGTGACCGGCTCAGACACACACAG 666
Qy 539 TGGACAGAGAGGGGCTGTGAGGCTGTGAGGCGCTGGGGGCGGGAGGCGCTTCCGCA 598
Db 667 TGGACAGAGAGGGGCTGTGAGGCTGTGAGGCGCTGGGGGCGGGAGGCGCTTCCGCA 726
Qy 599 CCGCAAGCTCAGCTCATGCCCTGTGTTCCAGGGAAGATCTCGGGCCAGAGAGCTGTCA 658
Db 727 CCGCAAGCTCAGCTCATGCCCTGTGTTCCAGGGAAGATCTCGGGCCAGAGAGCTGTCA 786
Qy 659 GCACTGCAGAGCTCCCGAGAGATGGGCTGTCTGTGTGCGGAGTCCGCGCCACCCGG 718
Db 787 GCACTGCAGAGCTCCCGAGAGATGGGCTGTCTGTGTGCGGAGTCCGCGCCACCCGG 846
Qy 719 GCACTGCAGAGCTCCCGAGAGATGGGCTGTCTGTGTGCGGAGTCCAGTACAGG-CTGTG 776
Db 847 GCACTGCAGAGCTCCCGAGAGATGGGCTGTCTGTGTGCGGAGTCCAGTACAGG-CTGTG 906
Qy 777 GCG-GACGGCCACCTCACAATGATG--AGGCGCTGTCTTCTGCAACCTCAT-GGACAT 832
Db 907 GCGGGAGGGCCACCTCACAATGATGAGAGGCGGTGTTCTTCTGCAAACTCATGGAGAT 966
Qy 833 GGTGA--GCATTACAGCAAGGAGGCGCTTCT 867
Db 967 GGTGNAGACATTACCGCAAGGAGGAGGCGCTAT 1003
RESULT 14
LOCUS BU617908 733 bp mRNA linear EST 23-SEP-2002
DEFINITION UI-H-DF0-bet-k-12-0-UI-s1 NCI-CGAP-DF0 Homo sapiens CDNA clone
UI-H-DF0-bet-k-12-0-UI 3', mRNA sequence.
ACCESSION BU617908
VERSION BU617908.1 GI:23284123
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 733)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gcraps-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, Bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA-yes.

FEATURES
source location/Qualifiers
1..733
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-DF0-bet-k-12-0-UI"
/tissue_type="Subchondral Bone"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI-CGAP_DF0"
/note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: Ecor I; Site_2: Not I;
NCI-CGAP_DF0 is a cDNA library containing the following
tissue(s): Subchondral Bone. The library was constructed
according to Bonaldo, Lennon and Soares' Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an Ecor I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is GTTAAAGCGTC.
TAG_L1B=UI-H-DF0
TAG_TISSUE=Subchondral bone
TAG_SEQ=GTAAAGCGTC"

BASE COUNT 100 a 248 c 226 g 158 t 1 others
ORIGIN

Query Match 35.4%; Score 707.6; DB 13; Length 733;
Best Local Similarity 99.2%; Pred. No. 8.8e-130;
Matches 721; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1274 CATGAGTACTTGGAGAGCAAGAAGCTTGTGACCGGACCTGCGCCCGCAACATCT 1333
|||||
DB 733 CATGAGTACTTGGAGAGCAAGAAGC-TGTGCACCGGACCTGCGCCCGCAACATCT 675
|||||

QY 1334 GGTCTCAGAGGAGCTGTGGCCCAAGTCAAGCACTTGTGCTGCGCAAGCCGAGCGAA 1393
|||||
DB 674 GGTCTCAGAGGAGCTGTGGCCCAAGTCAAGCACTTGTGCTGCGCAAGCCGAGCGAA 615
|||||

QY 1394 GGGGCTAGACTCAAGCCGGCTGCCGTCAAGTGAAGGCGCGCGAGGCTCTCAACAACG 1453
|||||
DB 614 GGGGCTAGACTCAAGCCGGCTGCCGTCAAGTGAAGGCGCGCGAGGCTCTCAACAACG 555
|||||

QY 1454 GAAGTTACCAAGCAAGTCGATGTCGTGAGTTTGGGCTGTGCTGCGAGGCTCTCTC 1513
|||||
DB 554 GAAGTTACCAAGCAAGTCGATGTCGTGAGTTTGGGCTGTGCTGCGAGGCTCTCTC 495
|||||

QY 1514 ATATGACAGGCTCTACCTTAATAATGTCACTGAAGAGGTGTGCGAGCGCGTGAGAA 1573
|||||
DB 494 ATATGACAGGCTCTACCTTAATAATGTCACTGAAGAGGTGTGCGAGCGCGTGAGAA 435
|||||

QY 1574 GGGGTACCGCATGGAACCCCGGAGGCTGTCAAGCCCGTGACGCTCTCATGAGCAG 1633
|||||
DB 434 GGGGTACCGCATGGAACCCCGGAGGCTGTCAAGCCCGTGACGCTCTCATGAGCAG 375
|||||

QY 1634 CTGCTGGAGGAGCAAGCCCGCGCCGACCTTCGCGAAATGTCGCGAGAGCTGGC 1693
|||||
DB 374 CTGCTGGAGGAGCAAGCCCGCGCCGACCTTCGCGAAATGTCGCGAGAGCTGGC 315
|||||

QY 1694 CCGGAGCTACGAGTGCAGGTGCGCCAGCCTCCGTCACAGGAGCGAGCGAGGCTC 1753
|||||
DB 314 CCGGAGCTACGAGTGCAGGTGCGCCAGCCTCCGTCACAGGAGCGAGCGAGGCTC 255
|||||

QY 1754 CACCTGCCCCGAAGCCAGAGAGCCCTGACCCACCCGCTGGGGCTTGGCCCCAGAGA 1813
|||||
DB 254 CACCTGCCCCGAAGCCAGAGAGCCCTGACCCACCCGCTGGGGCTTGGCCCCAGAGA 195
|||||

QY 1814 CCGAGAGATGAGAGTGGCGGCTGGGGGCACTGACAGGCGCCCAAGAGAGGTCCAGCGG 1873
|||||
DB 194 CCGAGAGATGAGAGTGGCGGCTGGGGGCACTGACAGGCGCCCAAGAGAGGTCCAGCGG 135
|||||

QY 1874 GCAGTCATCTCTGCTGGTGCACAGCAGGAGGCTGGGCCAGATGAGGGGGCTTGGGGCGC 1933
|||||
DB 134 GCAGTCATCTCTGCTGGTGCACAGCAGGAGGCTGGGCCAGATGAGGGGGCTTGGGGCGC 75
|||||

QY 1934 CCGTGAACACCCAGACCTGCGAAGATGATCGCCGATTAAGACGATTTCTAAGACTC 1993
|||||
DB 74 CCGTGAACACCCAGACCTGCGAAGATGATCGCCGATTAAGACGATTTCTAAGAAAA 15
|||||

QY 1994 TAAAAA 2000
|||||

DB 14 AAAAAA 8
|||||

RESULT 15
BX369553
LOCUS
DEFINITION BX369553 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
ACCESSION BX369553
VERSION BX369553.1 GI:30457823
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 986)
REFERENCE L1,W.B., Gruber,C., Jessee,J. and Polayres,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr; web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9238.r for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAF038ZF05.AF03591.2&cluster=9238.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradey Avenue genoscope sequence ID : CS0BAF038ZF05.AF03591_2.
Location/Qualifiers
1..986
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DU015Y001"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and Ecor V
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 207 a 270 c 321 g 187 t 1 others
ORIGIN

Query Match 35.4%; Score 707; DB 13; Length 986;
Best Local Similarity 95.0%; Pred. No. 1.3e-129;


```
Matches 828; Conservative 0; Mismatches 31; Indels 13; Gaps 9;
QY 840 CATTACGACGAGA-CAGGGCGCTATCTGCACCAAGCTGTGAGACAAAGCGAACA 898
    |||||
Db 3 CATTACGACGACGACCAAGGGCGCTATCTGCACCAAGCTGTGAGACCAAGCGAACA 62
QY 899 CGGGACCAAGTGGCCGAGAGAGAGCTGGCCAGGGCGGCTGGTTACTGAACCTGCAGCA 958
    |||||
Db 63 CGGGACCAAGTGGCCGAGAGAGAGCTGGCCAGGGCGGCTGGTTACTGAACCTGCAGCA 122
QY 959 TTTGACATTGGGAGCAGATGGAGAGAGAGTTGGAGCTGCTGCAGGGTGAGTA 1018
    |||||
Db 123 TTTGACATTGGGAGCAGATGGAGAGAGAGTTGGAGCTGCTGCAGGGTGAGTA 182
QY 1019 CCTGGGCAAAAGTGGCCGTGAAGATATCAAGTGTGATGTGACAGCCAGGCTTCCT 1078
    |||||
Db 183 CCTGGGCAAAAGTGGCCGTGAAGATATCAAGTGTGATGTGACAGCCAGGCTTCCT 242
QY 1079 GGAACGACGCGCCGCTATGACGAGATGCAACACGAAACTGTGCTCTCTGGCGT 1138
    |||||
Db 243 GGAACGACGCGCCGCTATGACGAGATGCAACACGAAACTGTGCTCTCTGGCGT 302
QY 1139 GATCCGACCAAGGGGCTGTACATGTCTAGAGCAGTGCAGCAAGGCAACCTGTTGAA 1198
    |||||
Db 303 GATCCGACCAAGGGGCTGTACATGTCTAGAGCAGTGCAGCAAGGCAACCTGTTGAA 362
QY 1199 CTTTCTGCGGACCCGGGGTCTGAGCCCTCGTGAACACCGCTCAGTCTCAGTTTCTCT 1258
    |||||
Db 363 CTTTCTGCGGACCCGGGGTCTGAGCCCTCGTGAACACCGCTCAGTCTCAGTTTCTCT 422
QY 1259 GCACGTGGCCGAGGGCATGAGTACCTGTGAGAGCAAGAGCTTGTGACCGGACCTGGC 1318
    |||||
Db 423 GCACGTGGCCGAGGGCATGAGTACCTGTGAGAGCAAGAGCTTGTGACCGGACCTGGC 482
QY 1319 CGCCCGCAACATCTGTCTCAGAGGACCTGGTGGCCAAAGTGCAGGCACTTGGCCTGGC 1378
    |||||
Db 483 CGCCCGCAACATCTGTCTCAGAGGACCTGGTGGCCAAAGTGCAGGCACTTGGCCTGGC 542
QY 1379 CAAAGCCGAGCGAAGGGGCTAGACTCAAGCCGGCTGCCCCGTCMACTGAGCGGCCGA 1438
    |||||
Db 543 CAAAGCCGAGCGAAGGGGCTAGACTCAAGCCGGCTGCCCCGTCMACTGAGCGGCCGA 602
QY 1439 GGCTCTCAACACAGGGAATTCAACGCAAGTCGGATGTCTGAG--TTTGGGGTGTGC 1497
    |||||
Db 603 GGCTCTCAACACAGGGAATTCAACGCAAGTCGGATGTCTGAGTTTGGGGTGTGC 662
QY 1498 TCTGGAGGCTTCTCATAT--GAGAGGGTCCGTAACCTTAAATGTCACTGAAGAGG 1554
    |||||
Db 663 TCTGGAGGCTTCTCATATGAGAGGGGCTCCGTAACCTTAAATGTCACTGAAGAGG 722
QY 1555 TGTGGAGGCC--GTGAGAGAGGGGTACCGCATGGAACCCCGAGGGC-TGTCCAGGCC 1611
    |||||
Db 723 TGTGGAGGCCCGGTGAGATAGGGTATACCGCATGGAACCCCGAGGGCTTGTCCAGGCC 782
QY 1612 CCGTGACGCTC-TATGAGCAGC-TGCTGGAGAGCAGA--GCCCGCCCGCGGCACCC 1667
    |||||
Db 783 CCGGGACCTTCTCATGAGCAGCTTGTGGGAAGCAGAACCTGCGCCCGGCCACCC 842
QY 1668 TTCCGCAACTGG-CCGAGAGACTGGCCCGG 1698
    |||||
Db 843 TTTCCGAAACTGTGTCGAGAAACTGTGCTCGG 874
```

Search completed: August 1, 2003, 19:41:12
Job time : 5081 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 15:26:49 ; Search time 590 Seconds
(without alignments)
6993.223 Million cell updates/sec

Title: US-09-977-260-1
Perfect score: 2000
Sequence: 1 ctgcctccaagtgttgacg.....attctaagcactcttaaaaa 2000

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2679534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PC1_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PC1US_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	2000	100.0	2000	9	US-09-977-269-1
2	2000	100.0	2000	10	US-09-977-260-1
3	2000	100.0	2000	11	US-09-977-261-1
4	1930	96.5	1989	14	US-10-084-817-341
5	547.6	27.4	2187	10	US-09-954-531-188
6	547.6	27.4	2187	15	US-10-298-377A-1
7	547.6	27.4	2420	14	US-10-177-293-87
8	192.6	9.6	2015	10	US-09-954-456-1983
9	192.6	9.6	2015	14	US-10-007-010-3
10	191.4	9.6	5527	10	US-09-880-107-3710
11	186.2	9.3	1911	10	US-09-917-800A-1611
12	185	9.2	1574	9	US-09-870-962-12
13	184.6	9.2	2354	10	US-09-967-768A-300
14	185.8	8.3	2451	10	US-09-771-161A-4
15	165.8	8.3	2674	13	US-10-003-295-1
16	162	8.1	2184	10	US-09-728-952-82

17	162	8.1	2520	11	US-09-953-047-3	Sequence 3, Appl1
18	160.4	8.0	3829	11	US-09-953-047-10	Sequence 10, Appl1
19	155.6	7.8	2298	14	US-10-175-523-50	Sequence 50, Appl1
20	148.8	7.4	2433	14	US-10-037-270-830	Sequence 830, Appl1
21	148.8	7.4	2454	10	US-09-964-824A-120	Sequence 120, Appl1
22	148.8	7.4	2598	14	US-10-072-036-110	Sequence 110, Appl1
23	148.8	7.4	2616	14	US-10-072-036-108	Sequence 108, Appl1
24	146.4	7.3	3454	13	US-10-044-090-48	Sequence 48, Appl1
25	146.4	7.3	3726	9	US-09-925-302-271	Sequence 271, Appl1
26	136.6	6.8	3945	9	US-09-921-771-4	Sequence 4, Appl1
27	136.6	6.8	3969	10	US-09-982-610-23	Sequence 23, Appl1
28	132.8	6.6	4698	10	US-09-895-652-1	Sequence 1, Appl1
29	131.6	6.6	3370	10	US-09-967-768A-144	Sequence 144, Appl1
30	127.2	6.4	2710	14	US-10-084-817-161	Sequence 161, Appl1
31	127.2	6.4	2820	10	US-09-924-859A-4	Sequence 4, Appl1
32	125.4	6.3	2301	14	US-10-242-943-3	Sequence 3, Appl1
33	124	6.2	4541	10	US-09-880-107-3785	Sequence 3785, Appl1
34	124	6.2	4541	14	US-10-123-036-3	Sequence 3, Appl1
35	123.4	6.2	3914	13	US-10-044-090-148	Sequence 148, Appl1
36	122	6.1	4221	14	US-10-220-801-8	Sequence 8, Appl1
37	122	6.1	6383	10	US-09-954-531-405	Sequence 405, Appl1
38	121.2	6.1	2574	9	US-09-735-103-2	Sequence 2, Appl1
39	121.2	6.1	2574	13	US-10-045-428A-2	Sequence 2, Appl1
40	121.2	6.1	5264	10	US-09-920-300A-1731	Sequence 1731, Appl1
41	121.2	6.1	5264	13	US-10-033-528-1731	Sequence 1731, Appl1
42	120.8	6.0	3633	9	US-09-725-433-1	Sequence 1, Appl1
43	120.8	6.0	10058	9	US-09-974-298-98	Sequence 98, Appl1
44	119.6	6.0	3416	8	US-08-987-689A-1	Sequence 1, Appl1
45	119.6	6.0	3416	14	US-10-292-524-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-977-269-1
Sequence 1, Appl1 Application US/09977269
Patent No. US20020082037A1
GENERAL INFORMATION:
APPLICANT: ULIRICH, AXEL
APPLICANT: GISHIZKY, MIKHAIL
APPLICANT: SURES, IRMINGARD
TITLE OF INVENTION: NOVEL MECAKARYOCYTIC PROTEIN TYROSINE KINASES
FILE REFERENCE: 038602/1260
CURRENT APPLICATION NUMBER: US/09/977,269
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 08/232,545
PRIOR FILING DATE: 1994-04-22
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 2000
TYPE: DNA
ORGANISM: Unknown Organism
FEATURE:
NAME/KEY: CDS
LOCATION: (258)..(1778)
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
US-09-977-269-1
Query Match 100.0%; Score 2000; DB 9; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CTCGCTCCAGTTGGACGCGGCGCTCGGGGTGACAGCCGCTCGGAGGCC 60
DB 1 CTCGCTCCAGTTGGACGCGGCGCTCGGGGTGACAGCCGCTCGGAGGCC 60
OY 61 TCTGTGGGGGCGGCGGCGGCTCGGGGCGGCGCTGAGCAGAGGAAGAAC 120
DB 61 TCTGTGGGGGCGGCGGCGGCTCGGGGCGGCGCTGAGCAGAGGAAGAAC 120

OY	121	AGGCGGGTCCAGTGGGACACCGAGCCCTACTCTCTGTGGCCAGCCGCTGGGCTGTGGCA	180
Db	121	AGGGCTGGGTCCAGTGGGACACCGAGCTCTCTACTCTCTGTGGCCAGCCGCTGGGCTGTGGCA	180
OY	181	GGCCATTCCACAGCGTCCCGGACGTGTGACACACTTGTGTCAATGTGCCTTCACACTGGCTAG	240
Db	181	GGCCATTCCACAGCGTCCCGGACGTGTGACACACTTGTGTCAATGTGTGCGCTTCACACTGGCTAG	240
OY	241	TTTCCCTTGGGGGGGCGATGGCGGGGGCGAGGCTCTGTGGTTTCTGTGGCGGCATTTCACG	300
Db	241	TTTCCCTTGGGGGGGCGATGGCGGGGGCGAGGCTCTGTGGTTTCTGTGGCGGCATTTCACG	300
OY	301	GCTGGATTTCAGTGAAGAACTTCCCGGGGTAGAGGCCCGCGTTCCTCCGAGCTGGGAC	360
Db	301	GCTGGATTTCAGTGAAGAACTTCCCGGGGTAGAGGCCCGCGTTCCTCCGAGCTGGGAC	360
OY	361	CCCCCTCCGCTCAGCCAGATGCCACAGAGGCGCTGGGGCCCGGGGACCCGAGTGTATCA	420
Db	361	CCCCCTCCGCTCAGCCAGATGCCACAGAGGCGCTGGGGCCCGGGGACCCGAGTGTATCA	420
OY	421	CCAAATGGGACACACCCCGCCCAAGCCAGGGGAGCTGGCTTCGCAAGGGCGACGTGG	480
Db	421	CCAAATGGGACACACCCCGCCCAAGCCAGGGGAGCTGGCTTCGCAAGGGCGACGTGG	480
OY	481	TCACCATTCTGGAGGCGCTGCGAGAAACAGACCTGGTTACCGGCTCAAGACACCAACACTTG	540
Db	481	TCACCATTCTGGAGGCGCTGCGAGAAACAGACCTGGTTACCGGCTCAAGACACCAACACTTG	540
OY	541	GACAGGAGGAGGCTGTGTGGCAGCTGGGGGCGTGGCGGAGCGGAGGCGCTTCGCGAGAC	600
Db	541	GACAGGAGGAGGCTGTGTGGCAGCTGGGGGCGTGGCGGAGCGGAGGCGCTTCGCGAGAC	600
OY	601	CCAAGCTCAGGCTCATGCGCTGTGTCCAGGGGAAAGATTTCGGGGCCAGGAGGCTGTCCAGC	660
Db	601	CCAAGCTCAGGCTCATGCGCTGTGTCCAGGGGAAAGATTTCGGGGCCAGGAGGCTGTCCAGC	660
OY	661	AGCTGCAAGCTTCCGAGAGATGGGCGTGTCTCTGTGTGCGGGAGTTCGCGCGCCACCCCGGCG	720
Db	661	AGCTGCAAGCTTCCGAGAGATGGGCGTGTCTCTGTGTGCGGGAGTTCGCGCGCCACCCCGGCG	720
OY	721	ACTACGTCCTGTGGCTGAGCTTTTGGCGCGCAGCTCATCTACCTACCGGCTGTGACCGCG	780
Db	721	ACTACGTCCTGTGGCTGAGCTTTTGGCGCGCAGCTCATCTACCTACCGGCTGTGACCGCG	780
OY	781	ACGGGCACTCAGCATGATGATGAGGCGCGTCTCTGTGAACTCATGAGCAATGGTGGAGC	840
Db	781	ACGGGCACTCAGCATGATGATGAGGCGCGTCTCTGTGAACTCATGAGCAATGGTGGAGC	840
OY	841	ATTACAGCAAGAGACAAAGGCGCTATCTGTGACCAAGCTGTGTAGAGCAAAAGGGAACACG	900
Db	841	ATTACAGCAAGAGACAAAGGCGCTATCTGTGACCAAGCTGTGTAGAGCAAAAGGGAACACG	900
OY	901	GGACCAAGTCGGCCGAGAGAGAGCTGGCCAGAGGGGGGCTGGTTACTAGCACTGCAACATT	960
Db	901	GGACCAAGTCGGCCGAGAGAGAGCTGGCCAGAGGGGGGCTGGTTACTAGCACTGCAACATT	960
OY	961	TGACATTGGGAGCAAGTCGGAAGGAGAGTTTGGAGCTGTCTCGTAGAGGTGATGCC	1020
Db	961	TGACATTGGGAGCAAGTCGGAAGGAGAGTTTGGAGCTGTCTCGTAGAGGTGATGCC	1020
OY	1021	TGGGGCAAAAAGTGGCCGTGAAGAAATATCAAGTGTGATGTGACAGGCCAGGCTTTCCTGG	1080
Db	1021	TGGGGCAAAAAGTGGCCGTGAAGAAATATCAAGTGTGATGTGACAGGCCAGGCTTTCCTGG	1080
OY	1081	ACGAGAGGCGCGTATGACGAAGATGTGCAACACGAGAACCTGTGTGCTCTCTGGCGCTGA	1140
Db	1081	ACGAGAGGCGCGTATGACGAAGATGTGCAACACGAGAACCTGTGTGCTCTCTGGCGCTGA	1140
OY	1141	TCTGTGACACAGGGGCTGTACTTTGTCTCATGTGAGCAGTGAGCAAGGGCAACCTGGTGA	1200
Db	1141	TCTGTGACACAGGGGCTGTACTTTGTCTCATGTGAGCAGTGAGCAAGGGCAACCTGGTGA	1200

QY	1201	TTCTCGGAGCCCGGGGTGAGCCCTCTGTGAACACGCTCAGCTCTGACAGTTTCTCTGC	1260
Db	1201	TTCTCGGAGCCCGGGGTGAGCCCTCTGTGAACACGCTCAGCTCTGACAGTTTCTCTGC	1260
QY	1261	ACGTGGCGGAGGGGCAATGAGTACTCTGGAGAGCAAGAAGCTTGTGACCCGGACCTGGCCG	1320
Db	1261	ACGTGGCGGAGGGGCAATGAGTACTCTGGAGAGCAAGAAGCTTGTGACCCGGACCTGGCCG	1320
QY	1321	CCCGCAACATCTGGTCTCTAGAGGACCTGTGGCCCAAGGTACGACCTTGGCTTGCCCA	1380
Db	1321	CCCGCAACATCTGGTCTCTAGAGGACCTGTGGCCCAAGGTACGACCTTGGCTTGCCCA	1380
QY	1381	AAGCGGAGGGGAAAGGGGCTAGACTCAAGCCGGGTGCGGTCAAGTGGAGCGGGCCGGAGG	1440
Db	1381	AAGCGGAGGGGAAAGGGGCTAGACTCAAGCCGGGTGCGGTCAAGTGGAGCGGGCCGGAGG	1440
QY	1441	CTCTCAACACACGGGAAGTTTACCACAGCAAGTCTGGATGTCTGAGTTTGGGGTCTCTCT	1500
Db	1441	CTCTCAACACACGGGAAGTTTACCACAGCAAGTCTGGATGTCTGAGTTTGGGGTCTCTCT	1500
QY	1501	GGGAGGTCTTCTCATATGAGACGGGCTCCGTACCTTAAATGTCACTGAAGAGGTGTGG	1560
Db	1501	GGGAGGTCTTCTCATATGAGACGGGCTCCGTACCTTAAATGTCACTGAAGAGGTGTGG	1560
QY	1561	AGGCGGTGGAGAAAGGGGTTACCGCATGGAAACCCCGGAGGGCTGTCCAGGCGCCGTGACG	1620
Db	1561	AGGCGGTGGAGAAAGGGGTTACCGCATGGAAACCCCGGAGGGCTGTCCAGGCGCCGTGACG	1620
QY	1621	TCCTCATAGCAGCTGTGTGGAGAGGACAGACCCGCGCGGCCACCTTCCGCAACTGG	1680
Db	1621	TCCTCATAGCAGCTGTGTGGAGAGGACAGACCCGCGCGGCCACCTTCCGCAACTGG	1680
QY	1681	CCGAGAACCTGGCCCGGGAGCTAGCGCAGTGCAGGTGCCCCAGCCTCCGTCTCAGGCGAGG	1740
Db	1681	CCGAGAACCTGGCCCGGGAGCTAGCGCAGTGCAGGTGCCCCAGCCTCCGTCTCAGGCGAGG	1740
QY	1741	ACGCCGAGGGCTTCACACTCTGGCCCCGAAAGCAGAGGCTCTGACCCACCCGGTGGGGCCCT	1800
Db	1741	ACGCCGAGGGCTTCACACTCTGGCCCCGAAAGCAGAGGCTCTGACCCACCCGGTGGGGCCCT	1800
QY	1801	TGGCCCCAGAGGACCGAGAGATGGAGATGTGGGCGCTGGGGGACCTGACACAGGCCCAAGG	1860
Db	1801	TGGCCCCAGAGGACCGAGAGATGGAGATGTGGGCGCTGGGGGACCTGACACAGGCCCAAGG	1860
QY	1861	AGGGTCAAGGGCGGCAAGTCATCTCTGTGGTGGCCCAAGCAGCAGGGGCTGGCCCACTAGGG	1920
Db	1861	AGGGTCAAGGGCGGCAAGTCATCTCTGTGGTGGCCCAAGCAGCAGGGGCTGGCCCACTAGGG	1920
QY	1921	GGCTTGGGGGGGGCCCGTGGACACCCCAAGACTTGGGAAGATGATGCCCGCATTAAGACGG	1980
Db	1921	GGCTTGGGGGGGGCCCGTGGACACCCCAAGACTTGGGAAGATGATGCCCGCATTAAGACGG	1980
QY	1981	ATTCTAAGAGACTCTAAAAA 2000	
Db	1981	ATTCTAAGAGACTCTAAAAA 2000	

SEQ ID NO 1
LENGTH: 2000
TYPE: DNA
ORGANISM: Unknown Organism
FEATURE:
NAME/KEY: CDS
LOCATION: (258)..(1778)
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
OTHER INFORMATION: kinase 1
US-09-977-260-1

Query Match 100.0%; Score 2000; DB 10; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGCTCCAGTTGTGCAAGCCGCGGAGCTTGGGGTGTGCAAGCGGCTCGCGAGAGCCC 60
DB 1 CTCGCTCCAGTTGTGCAAGCCGCGGAGCTTGGGGTGTGCAAGCGGCTCGCGAGAGCCC 60
QY 61 TCCTGG 120
DB 61 TCCTGG 120
QY 121 AGGCTCGGTTCAGTGGCAAGCCAGCTCCCTACCTCTGTGCGAGCCGCTGGCTGTGGCA 180
DB 121 AGGCTCGGTTCAGTGGCAAGCCAGCTCCCTACCTCTGTGCGAGCCGCTGGCTGTGGCA 180
QY 181 GGGCATTGCCAGAGTCCCGGAGCTGTGACACTTGTCTCACTGTGCTTCTACCTGCTCAG 240
DB 181 GGGCATTGCCAGAGTCCCGGAGCTGTGACACTTGTCTCACTGTGCTTCTACCTGCTCAG 240
QY 241 TTTCCCTGTGGGGGGGATGG 300
DB 241 TTTCCCTGTGGGGGGGATGG 300
QY 301 GCTGTGATTTCTGCTAGGAACTTCCCGGGGTAGCCCGCCGCTTCTCTCCAGCCTGGACCC 360
DB 301 GCTGTGATTTCTGCTAGGAACTTCCCGGGGTAGCCCGCCGCTTCTCTCCAGCCTGGACCC 360
QY 361 CCCCTCCGCTTCACCCAGAGATGCCAAGAGGCGCTGGGGCCCCGGGCAAGCTGTATCA 420
DB 361 CCCCTCCGCTTCACCCAGAGATGCCAAGAGGCGCTGGGGCCCCGGGCAAGCTGTATCA 420
QY 421 CCAATGGGAGACACCCGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGGGAGAGTGG 480
DB 421 CCAATGGGAGACACCCGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGGGAGAGTGG 480
QY 481 TCACCATCTGTGAGGCTGTGCGAACAAGAGCTGTACCGCTCAAGCAACCAACAGTGG 540
DB 481 TCACCATCTGTGAGGCTGTGCGAACAAGAGCTGTACCGCTCAAGCAACCAACAGTGG 540
QY 541 GACAGAGAGGGGCTGTGCGAGCTGGGGGGCTGCGGGAGCGGGAGCCCTCTCCGAGACC 600
DB 541 GACAGAGAGGGGCTGTGCGAGCTGGGGGGCTGCGGGAGCGGGAGCCCTCTCCGAGACC 600
QY 601 CCAAGCTGAGCTTCAATGCGGTGTTCCAGAGGAATCTCGGGCCAGAGAGGCTGTCCAC 660
DB 601 CCAAGCTGAGCTTCAATGCGGTGTTCCAGAGGAATCTCGGGCCAGAGAGGCTGTCCAC 660
QY 661 AGCTGAGCTTCCAGAGGATGGGCTGTCTGTGTCGGGATCCCGGCGCCACCCGGGGG 720
DB 661 AGCTGAGCTTCCAGAGGATGGGCTGTCTGTGTCGGGATCCCGGCGCCACCCGGGGG 720
QY 721 ACTAGTCTGTGCTGAGCTTTGGCCCGGAGCTATCCACTACCGCGTGTGCAACCGG 780
DB 721 ACTAGTCTGTGCTGAGCTTTGGCCCGGAGCTATCCACTACCGCGTGTGCAACCGG 780
QY 781 ACGGCACTTCAATTCATGAGGCGGTGTTCTTCTGCAACTCATGACATGATGGAGC 840
DB 781 ACGGCACTTCAATTCATGAGGCGGTGTTCTTCTGCAACTCATGACATGATGGAGC 840
QY 841 ATTACAGCAAGGACAAAGGGCGCTATCTGACCAAGCTGTGAGACCAAGGCGAAACAGC 900

DB 841 ATTACAGCAAGGACAAAGGGCGCTATCTGACCAAGCTGTGAGACCAAGGCGAAACAGC 900
QY 901 GGACCAAGTCCGCCAGAGAGAGCTGCGCAGGGCGGGCTGTACTGAACCTGCAACATT 960
DB 901 GGACCAAGTCCGCCAGAGAGAGCTGCGCAGGGCGGGCTGTACTGAACCTGCAACATT 960
QY 961 TGACATTGGGAGCACAGATCGAGAGGAGAGATTGGAGCTGCTCCGACGGGTGAGTACC 1020
DB 961 TGACATTGGGAGCACAGATCGAGAGGAGAGATTGGAGCTGCTCCGACGGGTGAGTACC 1020
QY 1021 TGGGGCAAAAGGTGGCCGTGAAGATATCAAGTGTGATGTGACAGCCAGGCTTCTCTGG 1080
DB 1021 TGGGGCAAAAGGTGGCCGTGAAGATATCAAGTGTGATGTGACAGCCAGGCTTCTCTGG 1080
QY 1081 ACGAGACGGCCCTCATGACGAAGATGCAACACGAGAACCTGGTGGCTCTCTGGGGCTGA 1140
DB 1081 ACGAGACGGCCCTCATGACGAAGATGCAACACGAGAACCTGGTGGCTCTCTGGGGCTGA 1140
QY 1141 TCCTGACACAGGGGCTGTACATTGTTCATGAGACAGTGTGACGAAGGCAACCTGTGAAC 1200
DB 1141 TCCTGACACAGGGGCTGTACATTGTTCATGAGACAGTGTGACGAAGGCAACCTGTGAAC 1200
QY 1201 TTTCTGCGACCCGGGGTTCAGGCCCTGTGTAACACCGCTACGCTCTGCAATTCTCTGTC 1260
DB 1201 TTTCTGCGACCCGGGGTTCAGGCCCTGTGTAACACCGCTACGCTCTGCAATTCTCTGTC 1260
QY 1261 ACGTGCGGAGGAGCATGAGATCTCTGAGAGCAAGAACTTTGCAACCGGCACTGGCCG 1320
DB 1261 ACGTGCGGAGGAGCATGAGATCTCTGAGAGCAAGAACTTTGCAACCGGCACTGGCCG 1320
QY 1321 CCCGCAATCTCTGCTGCTAGAGGACCTGTGAGGCAAGGTCAGAGCACTTGGCTGGGCA 1380
DB 1321 CCCGCAATCTCTGCTGCTAGAGGACCTGTGAGGCAAGGTCAGAGCACTTGGCTGGGCA 1380
QY 1381 AAGCCGAGGGAAGGGCTAGACTCAAGCCGGCTGCCCTCAAGTGTGAGCGGCCCGAGG 1440
DB 1381 AAGCCGAGGGAAGGGCTAGACTCAAGCCGGCTGCCCTCAAGTGTGAGCGGCCCGAGG 1440
QY 1441 CTCTCAAAACCGGGAAATTCAACAGCAAGTCCGATGTCTGAGATTTGGGGTGTCTCT 1500
DB 1441 CTCTCAAAACCGGGAAATTCAACAGCAAGTCCGATGTCTGAGATTTGGGGTGTCTCT 1500
QY 1501 GGGAGGCTTCTCATATGAGAGGGGCTCCGTACCTTAAATGTCACTGAAGAGGTGGG 1560
DB 1501 GGGAGGCTTCTCATATGAGAGGGGCTCCGTACCTTAAATGTCACTGAAGAGGTGGG 1560
QY 1561 AGGCGGTGAGAAGGGGTACCGCATGGAACCCCGAGGGCTGTCCAGGCCCGCTGCAGC 1620
DB 1561 AGGCGGTGAGAAGGGGTACCGCATGGAACCCCGAGGGCTGTCCAGGCCCGCTGCAGC 1620
QY 1621 TCTCATATGAGAGCTGTGAGGAGCAAGAGCCCGCCCGCGGCTCTCCGCAAACTGG 1680
DB 1621 TCTCATATGAGAGCTGTGAGGAGCAAGAGCCCGCCCGCGGCTCTCCGCAAACTGG 1680
QY 1681 CCGAGAAGCTGAGCCCGGAGCTACCGAGTGCAGAGTGCCTCCAGCTCTCGTCAAGGGCAG 1740
DB 1681 CCGAGAAGCTGAGCCCGGAGCTACCGAGTGCAGAGTGCCTCCAGCTCTCGTCAAGGGCAG 1740
QY 1741 ACGCCGAGGCTTCACCTTCGCCCCGAAGCCAGAGAGCCCTGACCCCAACCCGGTGGGGCCT 1800
DB 1741 ACGCCGAGGCTTCACCTTCGCCCCGAAGCCAGAGAGCCCTGACCCCAACCCGGTGGGGCCT 1800
QY 1801 TGGGCCGAGAGAGCCGAGAGTGTGAGAGTGTGGGGGCTGGGGGCACTGACAGGCCCAAG 1860
DB 1801 TGGGCCGAGAGAGCCGAGAGTGTGAGAGTGTGGGGGCTGGGGGCACTGACAGGCCCAAG 1860
QY 1861 AAGGTCACAGGCGGGAAGTCACTCTCTGTGCTCCACAGAGAGGCTGGCCACGTAAGG 1920
DB 1861 AAGGTCACAGGCGGGAAGTCACTCTCTGTGCTCCACAGAGAGGCTGGCCACGTAAGG 1920
QY 1921 GGGCTGTGGGGCGCCGTGAGCAACCCAGACCTGCAAGAGATGATGCCGATAAAGACG 1980

Dd	1921 GGCCTCTGGGCGGCCCTGGACACCCCGAGACTGCGAAGATGATCGGCCCATTAAGAAGCG	1980
Dy	1981 ATTCTAAGGACTCTAAAAA	2000
Dd	1981 ATTCTAAGGACTCTAAAAA	2000

```

RESULT 3
US-09-977-261-1
? Sequence 1, Application US/09977261
? Publication No. US20030054527A1
? GENERAL INFORMATION:
? APPLICANT: ULLRICH, AXEL
? APPLICANT: GISHIZKY, MIKHAIL
? APPLICANT: SURES, IRMINGARD
? TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
? FILE REFERENCE: 038602/1259
? CURRENT APPLICATION NUMBER: US/09/977,261
? CURRENT FILING DATE: 2001-10-16
? PRIOR APPLICATION NUMBER: 08/732,545
? PRIOR FILING DATE: 1994-04-22
? NUMBER OF SEQ ID NOS: 24
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 1
? LENGTH: 2000
? TYPE: DNA
? ORGANISM: Unknown Organism
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (258)..(1778)
? FEATURE:
? OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
? US-09-977-261-1

```

Query Match	100.0%	Score 2000;	DB 11;	Length 2000;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2000; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

OY	CTGCCTCAAGTTTGGACGCGGGGACCGCCTCGGGGTGTGGACGCGCGCTCGCGGAGGGCC	60
Db	1 CTGCCTCAAGTTTGGACGCGGGGACCGCCTCGGGGTGTGGACGCGCGCTCGCGGAGGGCC	60
OY	61 TCCTGGGGGGGGGCGCGGGGGGCGCTCGGGGGGCGGCCCTTGACAGAAAAACAGAAAGAAC	120
Db	61 TCCTGGGGGGGGGCGCGGGGGGCGCTCGGGGGGCGGCCCTTGACAGAAAAACAGAAAGAAC	120
OY	121 AGGTCGGGTCAAGTGGCAACCAAGTCCCTACCTCTGTGGCAAGCGCGCTGGCTGTGGCA	180
Db	121 AGGTCGGGTCAAGTGGCAACCAAGTCCCTACCTCTGTGGCAAGCGCGCTGGCTGTGGCA	180
OY	181 GGCCATTCCAGAGCTCCCGCGACTGTGACACATTGGTCAGTGGTCCTTCACCTGGCTGAG	240
Db	181 GGCCATTCCAGAGCTCCCGCGACTGTGACACATTGGTCAGTGGTCCTTCACCTGGCTGAG	240
OY	241 TTTCCCTCTGGGGGGGCGATGGCGGGGGGAGGCTCTGTGTTTCTGTGGCGGCAATTTCAG	300
Db	241 TTTCCCTCTGGGGGGGCGATGGCGGGGGGAGGCTCTGTGTTTCTGTGGCGGCAATTTCAG	300
OY	301 GCTGTGATTCTGCTGAGGAACTTCCCGGGGTGAGGCCCGCTTCTCCGAGCTGGCAAC	360
Db	301 GCTGTGATTCTGCTGAGGAACTTCCCGGGGTGAGGCCCGCTTCTCCGAGCTGGCAAC	360
OY	361 CCCCTCCCGTCTGCGAGGATGGCAAGAGAGGCGCTGGGGGCCCGGGGCAACCGAGTATACA	420
Db	361 CCCCTCCCGTCTGCGAGGATGGCAAGAGAGGCGCTGGGGGCCCGGGGCAACCGAGTATACA	420
OY	421 CCAAAATCGAGACACCGCCCAAGCGAGGGGAGTGGCCCTTCGCGAAGGGCGACGCTGG	480
Db	421 CCAAAATCGAGACACCGCCCAAGCGAGGGGAGTGGCCCTTCGCGAAGGGCGACGCTGG	480
OY	481 TCACATCTCTGGAGGCTGGGAGAACAGAGCTGTACCGGCTCAAGCACACACACGAGT	540
Db	481 TCACATCTCTGGAGGCTGGGAGAACAGAGCTGTACCGGCTCAAGCACACACACGAGT	540

D	b	481	TCACATCTCTGGAGAGGCTCTCGAGAACAAAGACTGTGTAACCGCTCAAGACCAACACAACTG	540
Q	y	541	GACAGGAGGGGGTCTGTGGACAGTGGGGGCTGCGGGAGCGGGAGGGCTCTCCGACAGC	600
D	b	541	GACAGGAGGGGGTCTGTGGAGCTGGGGGCTCTCGGGAGCGGGAGGGCTCTCCGACAGC	600
Q	y	601	CCAAAGCTCAACCTCATGCGCGTGGTTCCAGAGGGAAATCTCGGGCCAGGAGGCTGTCCAGC	660
D	b	601	CCAAAGCTCAACCTCATGCGCGTGGTTCCAGAGGGAAATCTCGGGCCAGGAGGCTGTCCAGC	660
Q	y	661	AGCTGCAGCCCTCCGAGAGATGGGGCTTTCTGTGCGGGATGCCGCGCCACCCGGAGC	720
D	b	661	AGCTGCAGCCCTCCGAGAGATGGGGCTTTCTGTGCGGGATGCCGCGCGCCACCCGGAGC	720
Q	y	721	ACTACGTCCTGTGAGCTTTTGGCCCGAGCGTATCCATCCAGCCGCTGTGCACCGC	780
D	b	721	ACTACGTCCTGTGAGCTTTTGGCCCGAGCGTATCCATCCAGCCGCTGTGCACCGC	780
Q	y	781	ACGGCCCACTCACAAATCGATGAGCGCGTCTTTCTGCACCTCATGCAATGTTGAGC	840
D	b	781	ACGGCCCACTCACAAATCGATGAGCGCGTCTTTCTGCACCTCATGCAATGTTGAGC	840
Q	y	841	ATTACAGCAAGGACAAAGGGCGGATGTGCACAACTGGTGGAGACCAAGGGGAAACAG	900
D	b	841	ATTACAGCAAGGACAAAGGGCGGATGTGCACAACTGGTGGAGACCAAGGGGAAACAG	900
Q	y	901	GGAACAAGTGGCGCGAGAGAGAGCTGGCCAGGGCGGGGTGTACTGAACCTGCAGCAT	960
D	b	901	GGAACAAGTGGCGCGAGAGAGAGCTGGCCAGGGCGGGGTGTACTGAACCTGCAGCAT	960
Q	y	961	TGACATTTGGGAGACACAGATCGGAGAGGAGATTTGAGCTGTCTCGCAGGTTGATACC	1020
D	b	961	TGACATTTGGGAGACACAGATCGGAGAGGAGATTTGAGCTGTCTCGCAGGTTGATACC	1020
Q	y	1021	TGGGGGAAAAAGTGGCGCGTGAAGAAATATCAATGATGTGACACCCAGGCTTTCCMG	1080
D	b	1021	TGGGGGAAAAAGTGGCGCGTGAAGAAATATCAATGATGTGACACCCAGGCTTTCCMG	1080
Q	y	1081	ACGAGACGGGCGTTCATGACGAAGATGCAACACGAGAACCTGGTGGCGTGA	1140
D	b	1081	ACGAGACGGGCGTTCATGACGAAGATGCAACACGAGAACCTGGTGGCGTGA	1140
Q	y	1141	TCCTGCACCAAGGGGCTGTACATTTGTCTATGAGACACTGTGACAAAGGGCAACTGGTGA	1200
D	b	1141	TCCTGCACCAAGGGGCTGTACATTTGTCTATGAGACACTGTGACAAAGGGCAACTGGTGA	1200
Q	y	1201	TTTCTGGGAGACCCGGGGTTCGAGCCCTCGTAAACACCGCTCAGTCCCTGCAAGTTTCTCGC	1260
D	b	1201	TTTCTGGGAGACCCGGGGTTCGAGCCCTCGTAAACACCGCTCAGTCCCTGCAAGTTTCTCGC	1260
Q	y	1261	ACGTGGCCGAGGGCATGAGATACCTGGAAGACAAAGACTTGTGCACCCGACACTGGCCG	1320
D	b	1261	ACGTGGCCGAGGGCATGAGATACCTGGAAGACAAAGACTTGTGCACCCGACACTGGCCG	1320
Q	y	1321	CCCGCAACATCTGTGTCATGAGAGACTGTGTGGCCAAAGTCAACGCACTTTTGGCTGGCA	1380
D	b	1321	CCCGCAACATCTGTGTCATGAGAGACTGTGTGGCCAAAGTCAACGCACTTTTGGCTGGCA	1380
Q	y	1381	AACCGAGGGGGAAGGGGGTGAAGCTAACCGCGGCTCCGCTAAAGTGGAGCGCGCCGAGG	1440
D	b	1381	AACCGAGGGGGAAGGGGGTGAAGCTAACCGCGGCTCCGCTAAAGTGGAGCGCGCCGAGG	1440
Q	y	1441	CTGTCAAAACACGGGAGTTCAACGACAAAGTCTGTGAGAGTTTGGGTGCTGTCT	1500
D	b	1441	CTGTCAAAACACGGGAGTTCAACGACAAAGTCTGTGAGAGTTTGGGTGCTGTCT	1500
Q	y	1501	GGGAGGTCTTTCATATGAGACGGGCTCCGTACCTTAAATGTCACTGAAGAGGTGTCCG	1560
D	b	1501	GGGAGGTCTTTCATATGAGACGGGCTCCGTACCTTAAATGTCACTGAAGAGGTGTCCG	1560
Q	y	1561	AGGCGCTGTGAAGAGGGGTACCCGATGGAACCCCGGAGGCTGTCCAGGCGCCGTGCAGC	1620
D	b	1561	AGGCGCTGTGAAGAGGGGTACCCGATGGAACCCCGGAGGCTGTCCAGGCGCCGTGCAGC	1620

Db	1260	GCACGTGGCCGAGGGGCAATGAGTACCTGAGAGAGCAAGAAAGTTTGTGGACACGGCAGCTGGC	1319
Qy	1319	CGCCCGCAACAATCTCGTCTCTAGAGAGACCTGTGTGCCCAAGTCAAGCCACTTTGGCCTGGC	1378
Db	1320	CGCCCGCAACAATCTCGTCTCTAGAGAGACCTGTGTGCCCAAGTCAAGCCACTTTGGCCTGGC	1379
Qy	1379	CAAGCCGAGGGGAGGGGCTAGACTCAAGCCGGCTGGCCCTCAAGTGTGAGCGGGGCCGA	1438
Db	1380	CAAGCCGAGGGGAGGGGCTAGACTCAAGCCGGCTGGCCCTCAAGTGTGAGCGGGGCCGA	1439
Qy	1439	GGCTCTCAACACGAGGAATTCACACAGCAATGCGAATCTGAGATTTTGGGTGTGCT	1498
Db	1440	GGCTCTCAACACGAGGAATTCACACAGCAATGCGAATCTGAGATTTTGGGTGTGCT	1499
Qy	1499	CTGGAGAGCTCTCATATGAGAGCGGCTCCGTACCTAAATGTACATGAAAGAGTGTCTC	1558
Db	1500	CTGGAGAGCTCTCATATGAGAGCGGCTCCGTACCTAAATGTACATGAAAGAGTGTCTC	1559
Qy	1559	GGAGGCGCTGAGAGAGGGGTACCGCATGGAACCCCGAGGGCTGTCCAGGCCCTGTGCA	1618
Db	1560	GGAGGCGCTGAGAGAGGGGTACCGCATGGAACCCCGAGGGCTGTCCAGGCCCTGTGCA	1619
Qy	1619	CGTCTCATATGACAGCTGTGTGGAGAGACAGCCCGCCCGCCGACCTTCCGCAACT	1678
Db	1620	CGTCTCATATGACAGCTGTGTGGAGAGACAGCCCGCCCGCCGACCTTCCGCAACT	1678
Qy	1679	GGCCGAGAACTGTGCCCCGGGAGACTACGAGTGTGACAGTGTCCGACCTCCGCTCAGGGCA	1738
Db	1679	GGCCGAGAACTGTGCCCCGGGAGACTACGAGTGTGACAGTGTCCGACCTCCGCTCAGGGCA	1738
Qy	1739	GGAGCGGAGAGCTGTGCCCCGGGAGACTACGAGTGTGACAGTGTCCGACCTCCGCTCAGGGC	1797
Db	1739	GGAGCGGAGAGCTGTGCCCCGGGAGACTACGAGTGTGACAGTGTCCGACCTCCGCTCAGGGC	1797
Qy	1798	CCTTGGCCCCAGAGAGACCGAGAGTGTGAGAGTGTGAGGCTGGGGGCACTGACAGGCCCA	1857
Db	1798	CCTTGGCCCCAGAGAGACCGAGAGTGTGAGAGTGTGAGGCTGGGGGCACTGACAGGCCCA	1857
Qy	1858	AGGAGGGTCCAGCGGGGCAATCATCTCTCGGTGGCCACAGCAGGGGCTGGCCACAGTA	1917
Db	1858	AGGAGGGTCCAGCGGGGCAATCATCTCTCGGTGGCCACAGCAGGGGCTGGCCACAGTA	1917
Qy	1918	GGGAGCTGTGGGGCGGCCCGTGAGACACCCAGACCTTGGGAAGAGATGTGCCCGATAAGA	1977
Db	1918	GGGAGCTGTGGGGCGGCCCGTGAGACACCCAGACCTTGGGAAGAGATGTGCCCGATAAGA	1977
Qy	1978	CGGATCTCAAG	1989
Db	1978	CGGATCTCAAG	1989

	Query Match	27.4%	Score 547.6	DB 10	Length 2187	
	Best Local Similarity	64.9%	Pred. No. 3.3e-12			
	Matches	846	Conservative	0	Mismatches	449
					Indels	9
					Gaps	2
QY	CTGGGCCCCGGGACCCAGTGTATCCACCAATGCGAGACACCCGCCCAAGCCAGGGA	454				
Db	CTGGGCATCCGGTACAGATGATATGCCAAGTACAACTTCCAGCGCACTGCCAGCAGGA	213				
QY	GCTGGCCCTCCGCAAGGGCGATGCTTACCATCTCGAGGCGCTTCGGAACAGACTG	514				
Db	CTCGCCCTTCGCAAAAGGAGCGTCTACCATTTGTGGCCGCTCACCAAGAGCCCACTG	273				
QY	GTACCGCGTCAGACACCACACAGTGGAGACAGAGGGGTGCTGGCAGCTGGGGCGCTCG	574				
Db	GTACAAAGCCA--AAACAGATGGGGCGGTGAGGGCATCATCCAGCCAACTACGTCA	330				
QY	GGAAGGGAGGAGCCCTCTCCGAGAGACCCCAAGCTCAGCTCATGCCGTTCACGGGA	634				
Db	GAACCGGAGGAGCGGAGGAGGCTGACCAAACTCAGCTCATGCTTGGTTCACGGGA	390				
QY	GATCGGGGCGAGGAGGCTGTCACAGACTGAGCTCCCGAGATGGGGCTGCTCGT	694				
Db	GATCACACGGAGCGAGGCTGAGCGCTTCTGTACCCGCGGAGACAGGCTGTTCTGCT	450				
QY	GCGGAGAGTCCGCGGCCACCCCGCGACTACGTCCTGTGAGCTTTGGCCGCGAGCT	754				
Db	GCGGAGAGACACAACTACCCCGGAGACTACACGCTGTGCGTGAAGCTCCGACGGCAAGT	510				
QY	CATCCACTACCGCGCTGTCACCGCGAGCGGCACTGCACCTCAGATTCAGAGCGCGTCT	814				
Db	GAGACTACCGCATCATGTATACCATGCGAGAGGCTGAGCATTCGACGAGGAGGTACTT	570				
QY	CTGCACACCTCATGAGCATGTGAGACATTACAGCAAGAGACAAAGGCGTATCTGCACCA	874				
Db	TGAGAACTCATGACAGCTGTGAGGACCTACACCTCAGACGAGATGAGACTGTGAA	630				
QY	GCTGCTGAGACCAAAAGCGGAAACACGGGACCAAGTCGGCCGAGAGAGAGCTGGCCAGGC	934				
Db	CCTCATTTAAACCAAAAGGTCATGAGGAGGAGACAGTGGCGGCCAGAGTGAAGTCTTACCGAG	690				
QY	GGCGTGTGTTAGAACTGACAGATTGATGACATTGGAGACAGATGCGAGAGGAGAGTT	994				
Db	CGGCTGGGCGCTGAACATGAAAGAGCTAAAGCTAGCTGTGACAGACATCGGGGAAGGGAGTT	750				
QY	TGAGAGCTGCTCGCAGAGGGTGAATACCTGCGGCAAAAAGGTGGCCGTGAAGAAATATCAAGTG	1054				
Db	CGGAGAGCTGATGTGGGGCGATTCACCGAGGGAACAAATGCCCGCTCAAGTCAATTAAGAA	810				
QY	TGATGTGACACCCAGGCTTCCTGAGAGAGAGCGGCGCTCATGATGAGAAATGACACAGA	1114				
Db	CGAGCGCACTGCCAGGCGCTTCTGAGCTGGAAGCTCAATGATGACGCAACTGCGGCAATG	870				
QY	GAACGTGTGCTGCTCTGCGGCGATGATCTCAGCCAG-----GGGCTGTACATGTGAT	1168				
Db	CACCTGTGTGAGCTCTCGGGCGATGATGTGAGAGAGAAAGGGCGGCTCTACATGTGAC	930				
QY	GGAGACGCTGAGCAAGGCAACCTGATGTAACCTTTCTGCGGACCGGGGTGCAAGCCCTGCT	1228				
Db	TGATGTACATGAGCAAGGGAGACCTGTGAGATACCTGTGGTCTAGGGTTCGATGATGCT	990				
QY	GAAACACCCCTCAGCTCCGCTGATTTTCTCTCAGCTGCGCAGGCGCATGAGATACCTGGA	1288				
Db	GAGGCGAGAGCTGTCTCTCAAGATTCGCTGATATGTCTGCGAGGCAATGGAATACCTGGA	1050				
QY	GAGCAGAAGCTTGTGACCGCGAGCTGGCGCCCGCAACATCTGTGCTCAGAGAGACTT	1348				

Db 1051 GGGCAACAATTTCTGTCAGACAGACCTGGCTGCCCGCAATGTCTGTCTGAGACAA 1110
Qy 1349 GGTGGCCAAAGTTCAGCACTTTGGCTGGCCAAAGCCGAGGAGGGCTAGACTCAAG 1408
Db 1111 GGTGGCCAAAGTTCAGCACTTTGGCTGGCCAAAGCCGAGGAGGGCTAGACTCAAG 1170
Qy 1409 CGGGTGGCCCTCAAGTGGAGGGCCCGGAGGCTTCAACACGGGAATTCACACGAA 1468
Db 1171 CAAGTGGCCAGTCAAGTGGAGGGCCCGGAGGCTTCAACACGGGAATTCACACGAA 1230
Qy 1469 GTGGATGTCTGAGATTTGGGGTCTCTGGAGAGTCTTCTCATATGACAGGGCTCC 1528
Db 1231 GTCTGAGCTGTGAGATTTTGGGAATCTTCTGGAATCTTCTTGGCCAGTGC 1290
Qy 1529 GTACCTTAAATGTCTACTGAAGAAGTGTCCGAGGCGGTGAGAGGGGTACCCGATGA 1588
Db 1291 TTATCCAAAGATTTCCCTGGAAGAGCTGTCCCTCGGTGGAGAAAGGCTACAGATGA 1350
Qy 1589 ACCCCCGAGGGCTGTCCAGGCCCCGTGACGCTCATATGAGCAGCTCTGGAGGAGA 1648
Db 1351 TGCCCCGAGGCTGCCCGCCGAGTCTATGAAGTCAATGAAGAACTCTGGACACTGGA 1410
Qy 1649 GCCCGCCCGGCGCCACCTTCCGAAACTGGCCGAGAGCTGG 1692
Db 1411 GCGCGCCATGCGGCGCTCTCTCTACAGCTCCGAGAGCAGCTTG 1454

RESULT 6

US-10-298-377A-1

Sequence 1, Application US/10298377A
Publication No. US20030130209A1
GENERAL INFORMATION:
APPLICANT: The Scripps Research Institute
APPLICANT: Cheresif, David A.
APPLICANT: Paul, Robert
APPLICANT: Elieceli, Brian
TITLE OF INVENTION: Method of Treatment of Myocardial
FILE REFERENCE: TSRI-651.5
CURRENT APPLICATION NUMBER: US/10/298,377A
CURRENT FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: 10/298,377
PRIOR FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: 09/470,881
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: 09/538,248
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: PCT/US99/11780
PRIOR FILING DATE: 1999-05-28
PRIOR APPLICATION NUMBER: 60/087,220
PRIOR FILING DATE: 1998-05-29
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2187
TYPE: DNA
ORGANISM: homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (134)...(1486)
US-10-298-377A-1

Query Match 27.4% Score 547.6; DB 15; Length 2187;
Best Local Similarity 64.9%; Pred No. 3.3e-121;
Matches 846; Conservative 0; Mismatches 449; Indels 9; Gaps 2;

Qy 395 CTGGGCCCCGGGACCCAGTGTATCCAAATGCGAGACACCCGCCCAAGCCAGGGA 454
Db 154 CTGGCATTCGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 213
Qy 455 GCTGGCTTCGCAAGGCGAGCTGTGATCCTCTGAGAGGCTTCGAGAACAGAGCTG 514
Db 214 CTTGCCCTTTCGCAAGGAGAGAGCTGTGATCCTCTGAGAGGCTTCGAGAACAGAGCTG 273

Qy 515 GTACCGCTCAAGACACACACAGTGTGACAGAGGAGGCTGTGCGAGCTGGGCGCTCG 574
Db 274 GTACCAAGCA-AAAACAAAGTGGGCGGTGAGGAGCTATCCAGGCACTACGTCCA 330
Qy 575 GGAGGGAGGAGGCTCTCCGAGACCCCAAGCTCAGCTCATGCGTGGTTCACAGGGA 634
Db 331 GAAGCGGAGGAGGCGTGAAGGCGGATGACCAACTCAGCTCATGCGTGGTTCACAGGGA 390
Qy 635 GATCTGGGCGGAGAGGCTGTCCAGACCTCAGCTCAGCTCCAGAGATGGGCTGTCTGG 694
Db 391 GATCAGGAGGAGGAGGCTGTGAGGCGCTTGTATCCCGGAGAGAGGCGCTGTCTGG 450
Qy 695 GCGGAGTCCCGGCGCCACCCCGGAGCTATGCTGCTGAGCTTGGCCGAGAGT 754
Db 451 GCGGAGGAGCAACAACTACCCCGGAGCTATGAGCTGTGCTGAGCTTCGAGAGAGT 510
Qy 755 CATCCATACCGGCTGTGACCGCGAGGCGCCACTCAATCGATGAGGCGGTGCTT 814
Db 511 GAGACATACCGCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
Qy 815 CTGCACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 874
Db 571 TGAGAACCTCATGAGCTGTGAGAGCACTACACCTCAGACGAGATGAGCTGTACGCG 630
Qy 875 GCTGTGAGACCAAGCGGAAACACGGGAGCCAGGAGGAGGAGCTGGCCAGGCG 934
Db 631 CCTCATTAACCAAGGATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 690
Qy 935 GGGCTGTACTGAACTGACAGCTTGTGAGATTTGGAGACAGAGATGAGAGGAGAGT 994
Db 691 GGGCTGTACTGAACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 750
Qy 995 TGAGCTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1054
Db 751 CGGAGAGCTGATGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 810
Qy 1055 TGATGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1114
Db 811 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 870
Qy 1115 GAACCTGTGCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1168
Db 871 CAACCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 930
Qy 1169 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1228
Db 931 TGAGTACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 990
Qy 1229 GAACAGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1288
Db 991 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1050
Qy 1289 GAGCAAGAAGCTGTGACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1348
Db 1051 GGGCAACAATTTCTGTCAGACAGACCTGGCTGCCCGCAATGTCTGTCTGAGACAA 1110
Qy 1349 GGTGGCCAAAGTTCAGCACTTTGGCTGGCCAAAGCCGAGGAGGGCTAGACTCAAG 1408
Db 1111 GGTGGCCAAAGTTCAGCACTTTGGCTGGCCAAAGCCGAGGAGGGCTAGACTCAAG 1170
Qy 1409 CGGGTGGCCCTCAAGTGGAGGGCCCGGAGGCTTCAACACGGGAATTCACACGAA 1468
Db 1171 CAAGTGGCCAGTCAAGTGGAGGGCCCGGAGGCTTCAACACGGGAATTCACACGAA 1230
Qy 1469 GTGGATGTCTGAGATTTGGGGTCTCTGGAGAGTCTTCTCATATGACAGGGCTCC 1528
Db 1231 GTCTGAGCTGTGAGATTTTGGGAATCTTCTGGAATCTTCTTGGCCAGTGC 1290
Qy 1529 GTACCTTAAATGTCTACTGAAGAAGTGTCCGAGGCGGTGAGAGGGGTACCCGATGA 1588
Db 1291 TTATCCAAAGATTTCCCTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1350

QY 1589 ACCCCCGAGGGCTGTCCAGGCCCGCCGAGCACTGCTCATGACGACGCTGCTGGAGGCGAGA 1648
Dd 1351 TGCCCGGAGGCGTGTGCCCGCCCGGAGCTATGAGAGCTCATGAAGACACTGCTGGACCTGGA 1410
QY 1649 GCCCGCCCGCGCGCCACCTTCCGCAAACTGCGCCGAGAGCTGG 1692
Dd 1411 CGCGCCCATGGGCGCCCTCTCTCTACACGCTTCGAGACCAAGCTTG 1454

RESULT 7

```

US-10-177-293-87
: Sequence 87, Application US/10177293
: Publication No. US20030124128A1
: GENERAL INFORMATION:
: APPLICANT: Lillie, James
: APPLICANT: Glatz, Karen
: APPLICANT: Zhao, Xunlei
: APPLICANT: Gannavarpu, Manjula
: APPLICANT: Kamatkar, Shubhang1
: APPLICANT: Mertens, Maureen
: APPLICANT: Myer, Vic
: APPLICANT: Wang, Youzhen
: APPLICANT: Xu, Yongyao
: APPLICANT: Hoersch, Sebastian
: APPLICANT: Monahan, John
: APPLICANT: Meyers, Rachel E.
: APPLICANT: Bast Jr., Robert C.
: APPLICANT: Hortobagyl, Gabriel N.
: APPLICANT: Puzsati, Lajos
: APPLICANT: Meric, Funda
: APPLICANT: Sahlin, Aysegul
: APPLICANT: Mills, Gordon B.
: TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT
: FILE REFERENCE: PREVENTION, AND THERAPY OF BREAST CANCER
: FILE REFERENCE: MRI-038
: CURRENT APPLICATION NUMBER: US/10/177,293
: CURRENT FILING DATE: 2002-06-21
: PRIOR APPLICATION NUMBER: US 60/299,887
: PRIOR FILING DATE: 2001-06-21
: PRIOR APPLICATION NUMBER: US 60/301,572
: PRIOR FILING DATE: 2001-06-27
: PRIOR APPLICATION NUMBER: US 60/306,501
: PRIOR FILING DATE: 2001-07-18
: PRIOR APPLICATION NUMBER: US 60/325,002
: PRIOR FILING DATE: 2001-09-25
: PRIOR APPLICATION NUMBER: US 60/362,585
: PRIOR FILING DATE: 2002-03-05
: PRIOR APPLICATION NUMBER: US 60/xxx,xxx
: PRIOR FILING DATE: 2002-05-14
: NUMBER OF SEQ ID NOS: 506
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 87
: LENGTH: 2420
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-177-293-87

```

Query Match	27.4%;	Score 547.6;	DB 14;	Length 2420;
Best Local Similarity	64.9%;	Pred. No. 3.3e-121;		
Matches 846;	Conservative 0;	Mismatches 449;	Indels 9;	Gaps 2

OY	395	CTGGGGCCCCGGGACCCAGTGTATATACCAAAATGGAGACACACCGGCCCAAGGCAGGGGA	454
Db	433	CTGGCCATCCGGTACAGAAATGTATTGGCCAGTACACTTCCACGGCACTGGCCGACGGA	492
OY	455	GCTGGCCTTCCGCAAGGGCGAGTGTACCATCTGTGAGGCGTCGCGAACAAGACTG	514
Db	493	CTGCCTCTCTCAAGGAGAGCTCTACCATTTGTGGCGTCCACGAAGACCCCAACTG	552
OY	515	GTAACCGCTCAAGCAACACAGTGGAGACAGAGGGCGTGTGGCACTGGGGCGCTGCG	574
Db	553	GTTCAAAAGCCA---AAAACAAGGTGGGCGTGAAGGGCATCTCCACGCAACTAAGTCCA	609

OY	575	GGACGGGAGGCGCCCTCCGACGAGACCCCAAGCACTGAGCTATGCGCGTGGTTCCAGGGGAA	634
Db	610	GAAGCGGGAGGGCGCTGAAGGCGGGTACCAAACTAGGCTATGCTTGTTCCAGGCGAA	669
OY	635	GATCTGGGGCCAGGAGGCTGTCAAGCAGACGTGCACCTCCGAGAGTGGGTGTTCCTCGT	694
Db	670	GATCAACAAGGAGAGAGGCTGAGGGCTTCTGTACCCCGGAGAGACAGGCTGTTCCGT	729
OY	695	GGGGAGTCCGCGGCCACCCCGGCGACTACGTCTGTGCGTGAAGCTTTGGCGCGAGT	754
Db	730	GCGGGAGAGCGACCAACTACCCCGGAGACTACACGCTGTGGTGAAGTGCACGCGCAAGT	789
OY	755	CATCCACTACCGGCTGTGACCCCGGACGGCGACACTCACAATGCATGAGGCCGTCTT	814
Db	790	GGAGCACTACCGCATCATGTACCTATGCGACGAAAGCTTACAGTACAGAGAGGTGATCTT	849
OY	815	CTGCACCTCTATGACATGTGTGAGCATTTACAGCAAGACAAAGGGCCCTATCTGCACCAA	874
Db	850	TGAACACTCATGTAGAGCTGTGGAGACACTACCTCAGACGACAGATGGACTGTACGGC	909
OY	875	GCTGTGAGACCAAAAGCGGAAACACGGGACCAGTGGCCGAGAGAGAGCTGGCCAGGC	934
Db	910	CCTCATTTAAACCAAAAGGTCAATGGAGGCGACAGTGGCGGCCAGAGATGATTTCTACCCAG	969
OY	935	GGGCTGGTCTTGAACCGGACGCTTTGACATTTGGAGCACAAGATGGAGAGGAGAGTT	994
Db	970	CGGCTGGGCCCTTGAAACTGAAGAGCGTAAGCTGTGACAGCCATTTCCGGAAGGGGAGTT	1029
OY	995	TGAGAGCTGCTCGAGGGGTGAGTACCTGCGGGGCAAAAGTGGCCGTGGAAGATATCAAGTG	1054
Db	1030	CGGAGAGGTGATGCTGGGGCATTTACGAGAGGAAACAAAGTGGCGTCAAGTGCAATTAGAA	1089
OY	1055	TGATGTGACAGCCAGGCTTCTGTGACAGAGAGCGCGCTCATGACGAATGCAACGAA	1114
Db	1090	CGAGCGCACTGCCAGGCGCTTCCGTGGAAGCCTCAGTATGACGCAACCTCGCGGCATAG	1149
OY	1115	GAACCTGGTGCTCCCGGGCGGATGCTCGACACAG-----GGGCTGAACATTTGAT	1168
Db	1150	CAACCTGATGAGTCTCTGGGCGGATGCTGTGAGAGGAAGGGCGGCTCTACATCTGCTAC	1209
OY	1169	GGAGCAGGTGAGCAAGGGCAACCTGCTGAACCTTCTCGGACCCGGGCTCGAGCCCTCGT	1228
Db	1210	TGATGTACATGGCCAAAGGGAGACCTTGTGAGACTACCTGGGCTTAGGGGTGGGTACGTGT	1269
OY	1229	GAACACGCTACGCTCGCATGATTTTCTGTGACAGTGGCCGAGGGCATGAGTACTGGA	1288
Db	1270	GGGCGGAGACTGTCTCTCAAGTCTCGTAGATGTCTCGAGGCGCATGGAATACCTGGA	1329
OY	1289	GAGCAAAAGCTTGTGACCGCGACCTGGCGCCCGCCGACAACTCTGGTCTCAGAGGACT	1348
Db	1330	GGGCAAAATTTCTGTGATCGAAGACCTGGCTGCCGCAAAATGTGTGTCTGTGAGGCAAA	1389
OY	1349	GGTGGCCAAAGTCAAGCACTTTGGCTCGGCGCAAAAGCCGAGCGGAAGGGCTTAGACTCAAG	1408
Db	1390	CGTGGCCAAAGTCAAGCACTTTGGTCTTCACCAAGAGGCTCCAGACCCAGACAGCGG	1449
OY	1409	CCGGCTGCCCTCAAGTGAACGCGCGCCGAGAGCTTCTCAAAACAGCGGAATTTCAACAGAA	1468
Db	1450	CAAGCTGCAGTCAAGTGAACAGAGCCCTGAGGCCCTGAGAGAGAAATTTCTCAACTAA	1509
OY	1469	GTCGGATGTCTGGAAGTTTGGGGTGTGCTCTGGAGAGTCTTCTCATATGACAGGGCTCC	1528
Db	1510	GTCGTGAGGTGTGGAGTTTGGAAATCTTCTCTGGGAATCTACTCTTTTGGCGAGAGGCC	1569
OY	1529	GTAACCTAAATATGCACTGAAAAGGTGTGGAGAGCCGTGGAACAAAGGGGTACCGCATGGA	1588
Db	1570	TTATTCCAAGAAATTCCTTGAAGAGCGTGTCCCTCGGGTGGAAAGGGCTTACAAAGATGA	1629
OY	1589	ACCGCCCGAGGAGGTGTGACAGGCCCGGCGACGTCCTATAGAGAGCTGTGGAGAGGAGA	1648
Db	1630	TGCCCCCGAGGCGCTGCCCGCCCGCATGTATGAAGTATATGAACACTGCTGTGGCACTGGA	1689
OY	1649	GCCCGGCCGCGGCGACCCCTTCGCAAAACTGGCCGAGAAAGCTGG	1692

Db 921 GTTTGGGAACTGTGATGGCCACTACAAAGACACCAAGGTGCAATGTAAGACAT 980
QY 1049 CAAG---TGATGATGACAGACCCAGGCTTCTGTGAGAGAGAGGCGTCATCGAAGAT 1105
Db 981 GAAGCAGGAGAGATGTCGTGTGAGAGGCTTCTGTGAGAGAGAGGCGTCATCGAAGAT 1040
QY 1106 GCACACAGAGACCTGTGTCGTCTGTGAGAGAGAGGCGTCATCGAAGAT 1165
Db 1041 GCAGCATGACAACTGTGCAAACTGTGAGAGAGAGGCGTCATCGAAGAT 1100
QY 1166 CAGGAGACAGTGTGAGAGAGGCGCACTGTGAGAGAGGCGTCATCGAAGAT 1225
Db 1101 CAGGAGATGATGAGAGAGAGAGAGGCGTCATCGAAGAGAGGCGTCATCGAAGAT 1160
QY 1226 CGTGAACAGGCTGTGAGAGAGGCGTCATCGAAGAGAGGCGTCATCGAAGAT 1285
Db 1161 GCAGCATGAGCAAACTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1220
QY 1286 GAGAGACAGAGAGCTGTGAGAGAGGCGCACTGTGAGAGAGGCGTCATCGAAGAT 1345
Db 1221 CGAGCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1280
QY 1346 CGAGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1393
Db 1281 CCGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1340
QY 1394 GGGGCTGAGACTCAAGGCGGCTGCGCTGCAAGTGTGAGAGAGGCGTCATCGAAGAT 1453
Db 1341 GCGTGTGAGAGAGGCGCAAGTGTGCGCTGCAAGTGTGAGAGAGGCGTCATCGAAGAT 1400
QY 1454 GAAGTGTGAGAGAGGCGCAAGTGTGCGCTGCAAGTGTGAGAGAGGCGTCATCGAAGAT 1513
Db 1401 CCGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1460
QY 1514 ATATGAGAGAGAGGCGCTGCGCTGCAAGTGTGAGAGAGGCGTCATCGAAGAT 1573
Db 1461 CCGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1520
QY 1574 GGGGCTGAGACTCAAGGCGGCTGCGCTGCAAGTGTGAGAGAGGCGTCATCGAAGAT 1633
Db 1521 TGGATGTGAGAGAGGCGCAAGTGTGCGCTGCAAGTGTGAGAGAGGCGTCATCGAAGAT 1580
QY 1634 CTGCTGTGAGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1670
Db 1581 CTGCTGTGAGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1617

RESULT 10
US-09-880-107-3710
Sequence 3710, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3710
LENGTH: 5527
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 X16416

US-09-880-107-3710
Query Match 9.6%; Score 191.4; DB 10; Length 5527;
Best Local Similarity 52.3%; Pred. No. 2,7e-36;
Matches 582; Conservative 0; Mismatches 501; Indels 30; Gaps 6;
QY 587 CCTCTCCGAGAGACCCAGAGCTGAGCTCATGCGGTGTTCACGAGAGATCTGGGCA 646
Db 492 CATCACGCAAGTCACAGCTGTGAGAAAGACTCTGTACATGAGCGCTGTCCGCA 551
QY 647 GAGAGCTGTCCAGACACTGAGCTCCGAGAGTGGGCTTCTGTCGGAGGCGCC 706
Db 552 TGCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 611
QY 707 GCGCCACCCCGGAG 766
Db 612 GAGCAGTCTGCGCAGAGGTCATGCGGTGAGATGAGAGAGAGAGAGAGAGAGAG 671
QY 767 ---CGTGTGACAGGAG 823
Db 672 GATCAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 731
QY 824 CATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 883
Db 732 GCGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 791
QY 884 ACCAAAGCGGAAACAGGAG 940
Db 792 TCGAGCCCAAG 851
QY 941 GTTATGAACTGTGAG 1000
Db 852 GAGATGAG 911
QY 1001 TGTCTGTGAG 1054
Db 912 GGTGTGAG 971
QY 1055 TGA---TGTGACAGCCAG 1111
Db 972 GAGACAGATGAG 1031
QY 1112 CGAGAGCTGTGAG 1168
Db 1032 CCTTAACGTGTGAG 1091
QY 1169 GAG 1228
Db 1092 TGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1151
QY 1229 GAACAGCTGTGAG 1288
Db 1152 GAAGAGCGTGTGAG 1211
QY 1289 GAGCAAG 1348
Db 1212 GAAGAAAG 1271
QY 1349 GGTGAG 1408
Db 1272 CTGTGTGAG 1331
QY 1409 CC-----GCGTCCCGTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1456
Db 1332 CCATGTGTGAG 1391
QY 1457 GTTACAG 1516
Db 1392 GTTCTCATGAG 1451
QY 1517 TGGAG 1576
Db 1452 TGGCATGTCCCTTACCGGAG 1511

FEATURE: -
OTHER INFORMATION: 507669
US-09-870-962-12

Query Match 9.2% Score 185; DB 9; Length 1574;
Best Local Similarity 56.0% Pred. NO. 8.7e-35;
Matches 378; Conservative 0; Mismatches 285; Indels 12; Gaps 1;

QY 1030 AGGTGGCCGTGAAGAATATCACTGATGTGACAGCCCGGACCTTCTGAGAGACGG 1089
DB 415 AGGAAGGCTTCATCCCTTCAATTTGTGCCAAGAACACCTGGAGCCCGAGGCCA 474
QY 1090 CCGTCTGACGAAGATGCAACACGACAACTGGTGGTCTCTGCGGCTGATCTTCACC 1149
DB 475 ACCTCATGAAGCAGCTGCAACACGACGCTGGTGGCTCTGCGTGGTGTGATCCACAG 534
QY 1150 AGGGGCTGTACATGTCATGAGACGAGTGAAGGCAACCGTGAATCTTCTGCGGA 1209
DB 535 AGCCCATCTACATCATCTGATGATGAGATGAGATGAGATGAGATGAGATGAGATG 594
QY 1210 CCGGCGGCTGAGCCCTGCTGACACCGCTGAGCTCTGAGTTTCTGACAGTGGCCG 1269
DB 595 CCGCTTACGACATCAAGTTGACCATCAACAACCTCGTGGACATGGACGCCAATTTGCA 654
QY 1270 AGGGCATGAGATGCTGAGAGACGACAACTGTGACCGGACCTGGCCCGGCAACA 1329
DB 655 AAGGCATGAGATGCTGAGAGACGACAACTGTGACCGGACCTGGCCCGGCAACA 714
QY 1330 TCGTGTCTGAGAGACGCTGAGTGGCCAGGTCAGCACTTGGCTGGC----- 1378
DB 715 TTCTGTGTCTGAGACCTGAGCTGACAGATGTCACACTTGGCTGAGCACTTGGCTG 774
QY 1379 -CAAGCCGAGCGGAGAGGGGCTGAGTCAAGCCGCTGCTGAGTGGAGCGCCCG 1437
DB 775 AGGACACAGGATGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 834
QY 1438 AGGCTCTCAACCGGGAAGTTTACCAAGAGTGGATGTCTGAGTGTGGGCTGCTGC 1497
DB 835 AAGCCATTTAACTGAGGAGATTCACCATCAAGTGTGATGTGGTGTGGATCTGC 894
QY 1498 TCGGAGGAGTCTCTCATATGAGAGGGGCTCCGTAACCTAAATGTACGAAAGAGTGT 1557
DB 895 TCGAGGAATTTGACACCGACGCGCATCTTACAGGAGGATGACCAACCGGAGGTGA 954
QY 1558 CGGAGGCGGTGAGGAAGGGGTACCGCATGGAACCCCGAGGCGTCTGACAGGCGCTGC 1617
DB 955 TTCAGAACCTGGAGGAGCTACCGCATGTGTGCGCTGACAACTGTCCAGAGAGACTGT 1014
QY 1618 ACCTCTCATGAGAGCTGTGGAGGACAGAGCCCGCGGCGGACCTTCCGCAAC 1677
DB 1015 ACCAATCTCATGAGGCTGTGTGAGAGAGAGCCGACAGAGAGAGAGAGAGAGAG 1074
QY 1678 TGGCGAGAGAGTGG 1692
DB 1075 TGGCGAGTGTGCTGG 1089

RESULT 13
US-09-967-768A-300
Sequence 300, Application US/09967768A
Patent No. US20020150877A1
GENERAL INFORMATION:
APPLICANT: Augustus, Meena
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
TITLE OF INVENTION: Sets
FILE REFERENCE: 689290-72
CURRENT APPLICATION NUMBER: US/09/967,768A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US/60/236,109
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,034
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,111

PRIOR FILING DATE: 2000-09-28
NUMBER OF SEQ ID NOS: 325
SOFTWARE: PatentIn version 3.0
SEQ ID NO 300
LENGTH: 2354
TYPE: DNA
ORGANISM: Homo sapiens
US-09-967-768A-300

Query Match 9.2% Score 184.6; DB 10; Length 2354;
Best Local Similarity 54.3% Pred. NO. 1.1e-34;
Matches 473; Conservative 0; Mismatches 374; Indels 24; Gaps 4;

QY 827 GGACATGTGAGACATTAACGAAGAGAGGCGCTATCTGACCAACCTGTGTGAGAC 886
DB 807 GAGAGTGTGACACATACATGAGAGTGAATGATGAGGAGGCTGTGCAACCTGTCTAT ----- 860
QY 887 AAGCGGAACACGAGCAACGATGCGCCGAGAGAGCTGTGCCAGGCGGCTGTACT 946
DB 861 CCGGCGCTGACCATCATGAGAGCGGAGAGAGCTGTGCCAGAGAGAGAGAGAGAGAG 920
QY 947 GAACTGTGACATTTGACATTTGGAGACACAGATGAGAGAGAGAGAGAGAGAGAGAG 1006
DB 921 CAGCGGACGCTCATGAGCTGAGAGCGCGCTGGGACCGGCTGCTTGGGAGTGTG 980
QY 1007 GCAGGCTGAGT---ACCTGGGCAAAAGGTGGCCGTGAAGATATCA---GTGTGATGT 1060
DB 981 GCTGGGACAGTGGAAAGCGGACGACATGAGTGGGCTGGAAGAGCTGTGAACCGGCGACCAT 1040
QY 1061 GACAGCCGAGCGCTTCTGAGAGAGAGCGCCGCTGATGACGAAGATGACAGAGAGAGCT 1120
DB 1041 GTCCCGGAGGCGCTTCTGAGAGAGAGCGGAGGATGATGAGTGTGCTGCGGACAGAGAGCT 1100
QY 1121 GGTGGCTGCTGAGGAGGAGTGTGATCTGACAGAGGAGGCTGATTTGATGATGAGAGAGT 1180
DB 1101 GGTGAGCTGTGAGCGCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1160
QY 1181 CAAGGCAACCTGTGAGAGCTTCTGAGAGAGCGGAGGAGAGAGAGAGAGAGAGAGAGT 1240
DB 1161 TCAAGGAGAGCTGTGATGATTTCTCAAGAACCGAGAGGAGAGAGAGAGAGAGAGAGT 1220
QY 1241 GCTGCTGAGATTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1300
DB 1221 ATTTGAGT 1280
QY 1301 TGGAGT 1360
DB 1281 CATTCAGT 1340
QY 1361 CAGCGAGTGTGAGT 1408
DB 1341 CCGAGAGTGTGAGT 1400
QY 1409 CCGGCTGCGGCTGAGT 1468
DB 1401 CAGGCTGCGGCTGAGT 1460
QY 1469 GTGAGATGTGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1528
DB 1461 GTGAGT 1520
QY 1529 GTACCTTAATGTCACTAAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1588
DB 1521 GTACCGAGT 1580
QY 1589 ACCCGCGGAGT 1648
DB 1581 GTGCGCTGAGT 1640
QY 1649 GCGCGCGGAGT 1679
DB 1641 CCGGAGT 1671

RESULT 14

US-09-771-161A-4

; Sequence 4, Application US/09771161A

; Patent No. US2002010811A1

; GENERAL INFORMATION:

; APPLICANT: LEVINE, et al.

; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES

; FILE REFERENCE: 802620-2005.1

; CURRENT APPLICATION NUMBER: US/09/771,161A

; CURRENT FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: 09/724,676

; PRIOR FILING DATE: 2000-11-28

; PRIOR APPLICATION NUMBER: 136776

; PRIOR FILING DATE: 2000-06-15

; PRIOR APPLICATION NUMBER: 135619

; PRIOR FILING DATE: 2000-04-12

; NUMBER OF SEQ ID NOS: 273

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 4

; LENGTH: 2451

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: -

; LOCATION: (1)..(2451)

; OTHER INFORMATION: "n" can be any nucleotide 'a', 'c', 'g' or 't'

US-09-771-161A-4

Query Match 8.3%; Score 165.8; DB 10; Length 2451;
Best Local Similarity 56.0%; Pred. No. 3.3e-30;

Matches 361; Conservative 0; Mismatches 272; Indels 12; Gaps 2;

QY	1060	TGACGCCAGGCGCTTCCTGAGCAGAGCGCGCTCATGCAAGATGCAACGAGAAC	1119
DB	1262	TGTCTCCAGAACCTTCCTGAGGAGGCCAACGTATGAGGCTCTGCGACGAGCGC	1321
QY	1120	TGTCGCTCTCTGCGGCGATGATCTGCAACAGGCGCTGATGTCATGTCATGAGCAGTGA	1179
DB	1322	TGTCGCACTGACGAGAGTGTGACCAAGAGCCCATCTACATGTCACCGAGTACATGG	1381
QY	1180	GCAAGGCAACCTGCTGAACTTCTGCGGACCCGGGCTGAGCCCTCGTGAACCCGCTC	1239
DB	1382	CCAGAGGATGCTGCTGATTTCTGAGACAGATGAAAGGAGAGGAGATTGTCATCCAA	1441
QY	1240	AGCTCTGAGTTTCTCTGACGTCGCGGCGAGGAGATGAGTACTGAGAGCAAGAAC	1299
DB	1442	GCTGATGATGATGTCGCGGAGATGCTGAAAGGATGATGATGATGATGATGATGAT	1501
QY	1300	TTGTCACCCGCGACCTGCGCGCCGCAACATCTGCTGCAAGAGACCTGTCGCGCAAG	1359
DB	1502	CCATCCACCGGAGCTGCGGCGGCGGCAACATCTGCTGATGAGGCTTTGCTGCAAAA	1561
QY	1360	TCAGGACTTTGGCTGGC-----CAAAAGCGAGCGAAGGGGCTAGACTCAAGCC	1410
DB	1562	TTGCTGATTTGGCTGGCTGCAATCATGACAGTGAATACAGGCCCAAGGGGGCCA	1621
QY	1411	GCTGCGCGTCAAGTGGAGCGGCGCGGAGGCTCTCAAAACGCGGAAGTTTCAACAGCAAT	1470
DB	1622	AGTTCCTCCATCAAGGAGACGCCCGGAAGCCATCCATTCGGGGCTTTCACATCAAAAG	1681
QY	1471	CGAGTCTGAGATTTGGGAGTGTCTGTCGAGGTCCTTCATATGAGCGGGCTCGT	1530
DB	1682	CAGAGTGTGTGCTTTTGGAGTCTCTGATGAGTGTCTCATTTATGGGCGGGTCCAT	1741
QY	1531	ACCTTAAATGTCTACTGAAAGAGTGTGCGAGCCGTGAGAGAGGAGTACCCCATGGAAC	1590
DB	1742	ACCCGAGGATGAGCAACCCCGAGGTCAATCCGCAACCTGGAGCGCGCTACCGCATGCCG	1801
QY	1591	CCCCGAGGAGTGTCTCAAGGCCCGT---CAGGTCTCATGAGCAGCTGTCTGGAGGAG	1647
DB	1802	GCCCGGACACTGCGCGCGCGAGGTGTACCGGCGGCTATCGCCGAGTGTGGCGCAGCC	1861

RESULT 15

US-10-003-295-1

; Sequence 1, Application US/10003295

; Publication No. US20020168741A1

; GENERAL INFORMATION:

; APPLICANT: GAN, Weiniu et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: C1001183DIV

; CURRENT APPLICATION NUMBER: US/10/003,295

; CURRENT FILING DATE: 2001-12-06

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 2674

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-003-295-1

Query Match 8.3%; Score 165.8; DB 13; Length 2674;
Best Local Similarity 55.6%; Pred. No. 3.4e-30;

Matches 370; Conservative 0; Mismatches 277; Indels 18; Gaps 2;

QY	1042	AGATATCATAGTGTATGTATGACAGCCGAGGCTTCTCTGAGAGAGAGGCGCTCATGACA	1101
DB	1642	AGAGCTCCACCTGACCTCAAGGCGCAAGTTTCTACAGGAAGCGAGATCTGAAAGCAGT	1701
QY	1102	AGATCAACAGAGAACTGTGTCCTCTCTGAGGCTATCTGCAACAGGAGGCTGTACA	1161
DB	1702	ACAGCCACCCCAACATCTGCTGCTCATTTGCTGTGACCCAGAGAGAGCCCATCTACA	1761
QY	1162	TTGTCATGAGACATGAGCAAGGCAACCTGATGAACTTCTGCGGAGCCGCGGCTCGAG	1221
DB	1762	TGCTATGATGAGTTTGTGAGGAGGCGGAGCTTCTTCTCTGCGACAGAGGAGGCGCC	1821
QY	1222	CCCTGCTGACACCGCTCAGCTCTGCAAGTTTCTCTGACGCTGCGGAGGAGGAGT	1281
DB	1822	GCCTGAGGTGAAGACTG---CTGCAAGTGTGGGAGATGCACTGCTGCGATGAGT	1878
QY	1282	ACCTGGAAGCAAGAGCTTGTGCAACCGGAGCTGCGCGCGCAACATCTCTGCTCAG	1341
DB	1879	ACCTGGAAGCAAGAGCTGCAATCCACGAGAGCTGCTGCGAACTGCTGCTGAGCAG	1938
QY	1342	AGGACCTGATGAGCAAGGTCAGGACCTTGGCTGCGCAAGAGCGAGGAGGAGCTAG	1401
DB	1939	AGAAAGATGTCTGTAAGATCACTGCTTTGGAGTATCCCGAGAGAGAGCCGATGGGTCT	1998
QY	1402	ACTCA-----AGCCGCTGCGCGCTGCAAGTGTGAGCGGCGCGAGGCTCTCA	1446
DB	1999	ATGAGCGCTCAGAGGCGCTCAGAGAAAGTCCCGTGAAGAGAGCCGACACTGAGCGCTTA	2058
QY	1447	AACAGGGAAGTTCAACAGCAAGTGTGATGTCTGAGTTTGGGCTGCTCTGGAGG	1506
DB	2059	ACTAGGCGGCTACTCTCCGAAAGCGAGCTGTGAGGCTTTGGCACTTCTGCTGGAGA	2118
QY	1507	TCTTCTCATATGAGGCGGCTCCGTAACCTTAAATGTCACTGAAAGAGGTGTGAGGCGG	1566
DB	2119	CTTTAGCGCTGGGCGCTTCCCTCTATCCCACTCAAGCAATCAGCAAGACAGGAGTTTG	2178
QY	1567	TGAGAGAGGATACCGCATGGAACCCCGAGAGGCTGTCAAGGCGCGCTGCACTCTCA	1626
DB	2179	TGAGAGAGGAGGCGCTGTGCGCTCCAGAGAGCTGTGCTGTGATCCGCTGAGGCTCA	2238
QY	1627	TGAGAGCTGCTGGAGGAGCAGAGCCGCGCGCGCAACCTTCCGCAAACTGGCGGAGA	1686
DB	2239	TGAGACAGTGTGCGGCTATGAGCTTGGGAGGCGCGCAGCTTCAAGCACTTCAACAGG	2298

Qy	1687	AGCTG	1691
Db	2299	AGCTG	2303

Search completed: August 1, 2003, 19:54:25
Job time : 597 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2003, 19:41:19 ; Search time 3240 Seconds
(without alignments)
3803.200 Million cell updates/sec

Title: US-09-977-260-2
Perfect score: 2671
Sequence: 1 MAGRSLVSMRAFHGDSAE.....PASYSGQADSTSPRSQEP 507

Scoring table:
BLOSUM62:
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 segs, 1215238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-O=/g9n2.1/USPTO_sppol/US99977260/runat_29072003_092741_14636/app.query.fasta.1.647
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09977260.ecgn.1.1.4237 -runat_29072003_092741_14636 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGGLOS
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlinu:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gp_estcl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estlinu:*
16: em_estlom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_dln:*
20: em_gss_vrl:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_pdg:*
26: em_gss_vrl:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2287.5	85.6	2046	11 AK013606	AK013606 Mus muscu
2	1978	74.1	2001	11 AK079816	AK079816 Mus muscu
3	1510.5	56.6	1043	12 B0054024	B0054024 AGENCOURT
4	1421	53.2	967	9 AL538511	AL538511
5	1360.5	50.9	936	13 B0184285	B0184285 AGENCOURT
6	1330	49.8	1035	9 AL580639	AL580639
7	1288.5	48.2	986	13 BX369553	BX369553 BX369553
8	1231	46.1	1040	9 AL558929	AL558929
9	1226.5	45.9	1129	12 BM450510	BM450510 AGENCOURT
10	1216	45.5	1032	9 AL558805	AL558805
11	1209	45.3	906	12 BG744770	BG744770 602722891
12	1207	45.2	1012	9 AL580543	AL580543
13	1193.5	44.7	822	10 BG395801	BG395801 602458975
14	1156	43.3	949	9 AL567349	AL567349
15	1139	42.6	837	12 B1912704	B1912704 603176356
16	1134	42.5	772	10 BG756878	BG756878 602713275
17	1127.5	42.2	708	12 BM715548	BM715548
18	1119.5	41.9	805	10 BE781641	BE781641
19	1116	41.8	1138	10 BE791829	BE791829 601586106
20	1093	40.9	953	12 B1488576	B1488576 603021104
21	1093	40.7	931	12 B1549778	B1549778 603194514
22	1087	40.4	862	12 B1768288	B1768288 603056458
23	1080	40.4	607	12 BE257861	BE257861 601109450
24	1078	40.4	1102	9 BM561134	BM561134 AGENCOURT
25	1074	40.2	729	12 AV702542	AV702542
26	1069	39.8	735	10 BE779216	BE779216 601464660
27	1064	39.0	1000	14 B7125568	B7125568
28	1064	38.7	950	12 BE795492	BE795492 601590164
29	1060.5	38.2	624	12 B1133187	B1133187
30	1019	36.9	631	12 BE254257	BE254257 601114671
31	997	37.3	1130	12 BM552996	BM552996 AGENCOURT
32	995.5	37.3	1130	12 B1819535	B1819535 603036746
33	990.5	37.1	912	12 AK037036	AK037036 Mus muscu
34	986.5	36.9	3064	11 BQ954082	BQ954082 AGENCOURT
35	985	36.5	913	13 BM918304	BM918304 AGENCOURT
36	976	36.5	952	10 BE791472	BE791472 601586006
37	973	36.4	1014	12 BM925586	BM925586 AGENCOURT
38	958.5	35.9	1014	12 BE257454	BE257454 601114403
39	956	35.8	591	10 AK018182	AK018182 Mus muscu
40	943	35.3	1075	11 B0848605	B0848605 AGENCOURT
41	937	35.1	824	13 CD363397	CD363397 UT-M-GLO-
42	927	34.7	569	14 BM476098	BM476098 AGENCOURT
43	927	34.7	800	12 BM701130	BM701130
44	920	34.4	1080	12	
45	913	34.2	552	12	

ALIGNMENTS

RESULT 1
LOCUS AK013606 2046 bp mRNA linear HTC 05-DEC-2002
DEFINITION Mus musculus adult male hippocampus cDNA, RIKEN full-length
tyrosine kinase, clone:2900029A15 product:megakaryocyte-associated
enriched library, full insert sequence.
ACCESSION AK013606
VERSION AK013606.1 GI:12851034
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

JOURNAL MEDLINE
PUBMED 99279253
REFERENCE 10349636
AUTHORS

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL MEDLINE
PUBMED 20499374
REFERENCE 11042159
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sunji, N., Ishii, Y., Nakamura, S., Hazama, M., Nishibe, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Motani, K.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multiplexed sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE
PUBMED 20530913
REFERENCE 11076861
AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Aizawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, T., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,
Saito, T., Okazaki, Y., Gojodori, T., Bono, H., Kasukawa, T., Saito, R.,
Kaoto, K., Matsuda, H., Ashburner, M., Batalov, S., Casavola, T.,
Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nakadoi, I., Pesole, G.,
Quackenbush, J., Schmitt, L. M., Stubbli, F., Suzuki, R., Tomita, M.,
Wagner, T., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bull, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hotman, M., Hume, D. A., Kamlay, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-Oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilmink, L.,
Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S.,
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

JOURNAL MEDLINE
PUBMED 21085660
REFERENCE 11217851
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL MEDLINE
PUBMED 12107466
REFERENCE 12107466
AUTHORS

6 (bases 1 to 2046)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arikawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nemura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toyo, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
Direct Submission

JOURNAL MEDLINE
PUBMED 12107466
REFERENCE 12107466
AUTHORS

7 Submitted (10-JUL-2000) yoshihide Hayashizaki. The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome

[illegible]

Db 208 CCCGGTAGCTAACCTTCGAAAGGCTGACATGGTGCACCATCTTGGAGGCCCTGTGAGGAC 267
 Qy 84 LysSerTrpTyrArgValLysHisHisThrSerGlyIngluGlyLeuLeuAlaGly 103
 Db 268 AAGAGCTGTATCGAAGCCAAACACCATGGCGGAGAGAGGCTGTCTGCGGCCCT 327
 Qy 104 AlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMetProTrpPhe 123
 Db 328 GCTGTGGACAGACGGGAGGCCCTCTCCACAGACCCCAAGCTCAGCTCATCTCATGTGTTT 387
 Qy 124 HisGlyLysIleSerGlyIngluAlaValGlnIngluLeuInProProGluAspGlyLeu 143
 Db 388 CATGGCAAGATCTCCGGCCAGGAAGCATACAGCAGCTGACGCCACCGAGAGGCGCTG 447
 Qy 144 PheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuGlyValSerPheGly 163
 Db 448 TTCCTGTGGAGGAAATCAGCTGCTCACCCTGGACACATCTGTGTGTGTGATTTGGC 507
 Qy 164 ArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIleAspGluAla 183
 Db 508 CGTACGCTACCTACCTACCGCTGTTGGATGAGATGGGACCTCAGCATGATGAGGCC 567
 Qy 184 ValPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLysAspLysGlyAlaIle 203
 Db 568 GTGTCTTCTGTAACTGTATGGACATGGTGGAGCAGTACACCAAGGACAAAGGGGCCATC 627
 Qy 204 CysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGluGluLeu 223
 Db 628 TGCACCAAGCTGTGTGAAGCCAGAGGACGAGGAGGCGCAAGTGTGCGAGAGGAGAGCTC 687
 Qy 224 AlaArgAlaGlyTyrPleuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnIleGlyGlu 243
 Db 688 GCCAAGCTGTGCTGACTGCTGACTGCTGACTGCTGACTGCTGAGAGCAGATTTGGAGAG 747
 Qy 244 GlyGluPheGlyAlaValLeuGlnGlyLysTyrLeuGlyGlnLysValAlaValLysAsn 263
 Db 748 GGGAGTTTGGAGCGCTGCTACAGGCTGAGTACCTGGAGCAGGAGGCTGCTGAGAGAT 807
 Qy 264 IleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMetThrLysMet 283
 Db 808 ATCAAGTGTATGTGTGACAGCCGAGCCCTCTCTGATGAGAGCGCTGTGATACCAAGCTG 867
 Qy 284 GlnHisGluLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeuTyrIleVal 303
 Db 868 CAGCAGAGACCTAGTCCGACTCCTGGGTGTGATCTGCAACCAAGGCTGTACATGTCTC 927
 Qy 304 MetLysValSerLysGlyAsnLeuValAsnPheLeuAlaGlyThrArgGlyArgAlaLeu 323
 Db 928 ATGGAGCAGCTGTGACCAAGGCGAACCTGTGTAACTCTGCGCCAGCGGGGCCGTGCTCTT 987
 Qy 324 ValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlnGlyMetGluTyrLeu 343
 Db 988 GTGAGCAGCTCTCAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1047
 Qy 344 GluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuValSerGluAsp 363
 Db 1048 GAGAGCAAGAGCTGTGCTCAGCCGAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1107
 Qy 364 LeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGlyLeuAspSer 383
 Db 1108 TTGGTGTGGCCAGGCTGAGCTGTGGCTGTACCAAGGAGAGGCCCAAGGGCTGTGAGCTCA 1167
 Qy 384 SerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLysPheThrSer 403
 Db 1168 AGCGGCTGTGAGCTGAGCTGAGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1227
 Qy 404 LysSerAspValTyrPheSerPheGlyValLeuLeuTyrPgluValPheSerTyrGlyArgAla 423
 Db 1228 AAGTCGATGCTGTGAGTTGGGGTGTGTGTGGGAAGCTCTCTTATGGAAGAGACC 1287
 Qy 424 ProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGlyTyrArgMet 443
 Db 1288 CCATACCCCAAGATGTGCTTAAAGAGGTTTCAAGAGGCTGTGGAGAGGTTTACCGCATG 1347

Qy 444 GluProGluGluGlyCysProGlyProValHisValLeuMetSerSerCysTrpGluAla 463
 Db 1348 GAGCCCCCATGAGCTGCCAGGCTGTGTGACACACCTCATGTGGTACTGCTGGAGAGCA 1407
 Qy 464 GluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArgGluLeuArg 483
 Db 1408 GAGCTGTGCGGCCGAGCACCCTTCCGCAAAATAGTGAAGGCTGAGGCGGCTGAGCTCCG 1467
 Qy 484 SerIleGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThrSerProArg 503
 Db 1468 AGTGTGGGTGTCTGCGCCCGCTGCGGGGACAGAGGCTGAGGCTGAGCTCCACACAGG 1527
 Qy 504 SerGlnGluPro 507
 Db 1528 AGCCAGGAGCCCC 1539
 RESULT 2
 AK079816
 LOCUS
 DEFINITION
 Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched library, clone:430076c03 product:megakaryocyte-associated tyrosine kinase, full insert sequence.
 ACCESSION
 AK079816
 VERSION
 AK079816.1 GI:26098774
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 JOURNAL
 MEDLINE
 PUBMED
 99279253
 10349636
 2
 REFERENCE
 3
 TITLE
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, N., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL
 MEDLINE
 PUBMED
 20499374
 11042159
 4
 REFERENCE
 5
 TITLE
 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, O., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matshiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichannel sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 JOURNAL
 MEDLINE
 PUBMED
 20530913
 11076861
 6
 REFERENCE
 7
 TITLE
 Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuhl, P., Lewis, S., Matsuo, I., Nikaido, I., Pesole, G., Quackenbush, J., Schiml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hotmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,

QY 222 GtLeuAlaArgAlaGlyTrrPheLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnIle 241
|||||
Db 543 GACCTGGCCAGGGCGGGCTGGTTACTGAACCTCCACACATTGTGACATTGGGAGCAGATC 602
QY 242 GtGlyGtGlyGluPheGlyAlaValLeuGlnGtGlyTyrLeuGlyGlnValAlaVal 261
|||||
Db 603 GAGAGGAGAGAGTTTGGAGCTGCTCCAGAGGTGATACCTGGGCAAAAGGTGGCGGTG 662
QY 262 LysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspIleThrAlaValMetThr 281
|||||
Db 663 AAGAAATATCAAGTGTATGTGACAGCCAGGCTCTCTGGAGAGAGCGGCCGTGATGAG 722
QY 282 LysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeuTyr 301
|||||
Db 723 AAGATCAACACAGAGAACCTGGTGGCTCTCCAGGCGGTATCTCCACACAGGCGGTGA 782
QY 301 rIleValMetGlnHisValSerLysGlnLysLeuValAsnPheLeuArgThrArg-Gly 320
|||||
Db 783 CATGTGATGAGACAGTGTGAGCAAGGCAACCTGGGTGAATCTTCTGGCGAGCCGGGGT 842
QY 321 ArgAla-LeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAla-GluGly 340
|||||
Db 843 CGAGCCCTCTGTGAACACCGCTCAGCTCCTGAGTTTCTCTGACCTGGCCCAAGCA 902
QY 340 erglu-TyrLeuGluSer---LysLysLeuValHisArgAspLeuAla 355
|||||
Db 903 TGGAAATACCTGTGAAGACCAAAAGCTTTGTGCACCCCGAAGCTGGCGCC 953

RESULT 4
ALS38511 967 bp mRNA linear EST 31-MAY-2003
LOCUS CS0DF027YD13 5-PRIME, mRNA sequence.
DEFINITION ALS38511
ACCESSION AL538511
VERSION AL538511.2 GI:31263101
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 967)
L1,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12802004.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191191006 EVRY cedex - France
Email: sequef@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9238.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DF027CB07OP1a;cluster=9238.r. Contact :
Feng Liang Email: fliang@life.technet.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS0DF027CB07OP1.
Location/Qualifiers
1. 967
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF027YD13"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone.lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain, Vector: pCMVSPORT.6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

BASE COUNT 193 a 297 c 330 g 147 t
ORIGIN

Alignment Scores:
Pred. No.: 1,78e-120 Length: 967
Score: 1421.00 Matches: 270
Percent Similarity: 94.10% Conservative: 1
Best Local Similarity: 93.75% Mismatches: 5
Query Match: 53.20% Indels: 12
DB: 9 Gaps: 1
US-09-977-260-2 (1-507) x ALS38511 (1-967)
QY 2 AlaGlyArgGlySerLeuValSerThrArgAlaPheHisGlyCysAspSerAlaGlu 21
|||||
Db 139 GCGGGCGCGGC-----TCGGGGCGGCC 162
QY 22 LeuProArgValSerProArgPheLeuArgAlaThrPHisProProValSerAlaArg 41
|||||
Db 163 CCGAGCAGGTGAGCCCGCTCTCCAGAGCTGGCAGCCCTCCCGCTGACGACG 222
QY 42 MetProThrArgArgThrAlaProGlyThrGlnCysIleThrLysCysGlnHisThrArg 61
|||||
Db 223 ATGCCAAGAGCGCGCTGGCGCCCGGCGACCCAGTGTATACCAATGCGAGCACACCCG 282
QY 62 ProLysProGlyGluLeuAlaPheArgGlyAspValValThrIleLeuGluAlaCys 81
|||||
Db 283 CCCAAGCCAGGGAGCTGGCTCTCCAGAGGCGAGCTGTCACCAATCTGGAGGCTGC 342
QY 82 GluAsnLysSerTrpTyrArgValLysHisHisThrSerGlyGlnGlyLeuAla 101
|||||
Db 343 GAGAACAGAGCTGTGATACCCGCTCAAGCACACACAGTGTGAGAGAGGCGCTGCGCA 402
QY 102 AlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMetPro 121
|||||
Db 403 GCTGGGGCGCTCGGAGCGGAGCGGCGCTCTCCGAGAGCCCAAGCTCAGCTCATGCGG 462
QY 122 TrpPheHisGlyLysLysLeuSerGlyGlnGluAlaValGlnGlnLeuGlnProProGluAsp 141
|||||
Db 463 TGGTTCCAGCGGAGATCTCGGGCCAGAGGCTGTCCAGACCTGACCTCCCGAGAGAT 522
QY 142 GlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuCysValSer 161
|||||
Db 523 GGGCTGTCTCTGGTGGGAGATCCGGCGCCAGCCCGGAGACACTCTCTGTCGTGAGC 582
QY 162 PheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIleAsp 181
|||||
Db 583 TTGGCGCGGAGCTGATCCACTACCGCGCTGTCCAGCGGAGCGGACCTCACAAATCGAT 642
QY 182 GluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLysAspLysGly 201
|||||
Db 643 GAGGCCGTGTCTTGTCAACCTCATGAGCATGTGTGAGCATTAACAGCAAGCAAGGCG 702
QY 202 AlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGluGly 221
|||||
Db 703 GCATGTGCACCAAGCTGTGAGAGCAACCAAGCGGAAACAGCGGACCAAGTCCGGCGAGAG 762
QY 222 GtLeuAlaArgAlaGlyTrrPheLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnIle 241
|||||
Db 763 GAGCTGGCCAGCGCGGGCTGGTTACTGAACCTGACACATTTGACATTGGGAGCAGATC 822
QY 242 GtGlyGtGlyGluPheGlyAlaValLeuGlnGtGlyTyrLeuGlyGlnValAlaVal 261
|||||
Db 823 GAGAGGAGAGAGTTTGGAGCTGCTCCAGAGGTGATACCTGGGCAAAAGGTGGCGGTG 882
QY 262 LysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspIleThrAlaValMetThr 281
|||||
Db 883 AAGAAATATCAAGTGTATGTGACAGCCAGGCTCTCTGGAGAGAGCGGCCGTGATGAG 942
QY 282 LysMetGlnHisGluAsnLeuVal 289
|||||
Db 943 AAGATCAACACAGAGAACCTGGTG 966

RESULT 5
BUI84285

LOCUS BU184285 936 bp mRNA linear EST 04-SEP-2002
DEFINITION AGENCOURT 7942749 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6171643
5', mRNA sequence.
ACCESSION BU184285
VERSION BU184285.1 GI:22698269
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 936)
NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM13540 Row: e Column: 20
High quality sequence stop: 674.
Location/Qualifiers
1..936
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6171643"
/tissue_type="retinoblastoma"
/lab_host="NIH_MGC_67"
/note="Organ: eye; Vector: PCMV-SPORE6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 184 a 268 c 314 g 169 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 6,35e-115 Length: 936
Score: 1360.50 Matches: 288
Percent Similarity: 77.17% Conservative: 6
Best Local Similarity: 75.59% Mismatches: 11
Query Match: 50.94% Indels: 77
DB: 13 Gaps: 3
US-09-977-260-2 (1-507) x BU184285 (1-936)
OY 99 LeuLeuAlaAlaGlyAlaLeuArgLUAArgLUAAlaLeuSerAlaAspProLysLeuSer 118
DB 11 GTGCTGGACGCTGGGGCGCTCCGGAGCGGAGCCCTCTCCGACGCCAACCTCAGC 70
OY 119 LeuMetProTrpPheHisGlyLysIleSerGlyGlnGlnAlaValaGlnGlnLeuGlnPro 138
DB 71 CTCATGCCGTGGTCCACGGGAAGATCTCGGGCCAGAGAGCTGTCCAGCACTGCAAGCTT 130
OY 139 ProGluAspGlyLeuPheLeuValaArgLUAuSerAlaArgHisProGlyAspTyrValLeu 158
DB 131 CCCGAGATGGGCTGTCTCTGCTGGAGAGTCCGCGCCACCCCGGGGAGCTACGCTCTG 190
OY 159 CysValSerPheGlyArgAspValaIleHisTyrArgValaLeuHisArgAspGlyHisLeu 178
DB 191 TGCCTGAGCTTTGGCCGGACGATCATCAACCGCTGTCGACCCGAGGCGCACCTC 250
OY 179 ThrIleAspGlnAlaValaPhePheCysAsnLeuMetAspMetValaGlnHisTyrSerLys 198
DB 251 ACATTCATGAGGCGGTCTTCTGCAACTCATGAGCATGTGTGGAGCATTTACAGCAAG 310
OY 199 AspLysGlyAlaIleCysThrLysLeuValaArgProLysArgLysHisGlyThrLysSer 218
|||||

DB 311 GACAAGGGCGCTATCTGCACCAAGCTGTGTGAGCAAAAGCGGAACAGGGAGCAAGTCTG 370
OY 219 AlaGluGluGluLeuAlaAlaArgAlaGlyTyrPheLeuAsnLeuGlnHisLeuThrLeuGly 238
DB 371 GCCGAGGAGGAGCTGGCGCAGGGCGGTGTTACTGAACCTGCAGCATTTGACATTGGGA 430
OY 239 AlaGlnIleGlyGluGluGluPheGlyValaValaLeuGlnGlyGlyTyrLeuGlyGlnLys 258
DB 431 GCACAGATCGGAGAGGAGAGATTGGAGCTCTCTGCAAGGTGAGTACTCTGGGCAAAAG 490
OY 259 ValAlaValaLysAsnIleLysCysAspValaThrAlaGlnAlaPheLeuAspGlyThrAla 278
DB 491 GTGGCCGTAAGATATCAAGTGTGATGTGACACCCGAGGCTCTCTGACAGAGAGGCC 550
OY 279 ValMetThrLysMetGlnHisGlnAsnLeuValaArgLeuLeuGlyValaIleLeuHisGln 298
DB 551 GTCATGACGAGAGTACCAACAGAAACCTGTGGCTCTCCGCGGCGTATCTCAGCAG 610
OY 299 GlyLeuTyrIleValaMetGlnHisValSerLysGlyAsnLeuValaAsnPheLeuArgThr 318
DB 611 GGGCTGTACATTTGTATGAGACAC-----GGCAACCTGTGMACTTTCTCGGAGCC 661
OY 319 ArgGlyArgAlaLeuValaAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValaAlaGlu 338
DB 662 CGGGGTGAGCCCTCGTGAACACCGCTCAGCTCTGCGAGTTTCTCT----- 708
OY 339 GlyMetGlyTyrLeuGluSerLysLysLeuValaHisArgAspLeuAlaAlaArgAsnIle 358
DB 708 ----- 708
OY 359 LeuValSerGluAspLeuValaIleAlaLysValSerAspPheGlyLeuAlaIleAlaGluArg 378
DB 708 ----- 708
OY 379 LysGlyLeuAspSerSerArgLeuProValaLysTrpThrAlaProGluAlaLeuLysHis 398
DB 708 ----- 708
OY 399 GlyLysPheThrSerLysSerAspValaTyrSerPhe--GlyValLeuLeuTrp--GluVal 417
DB 709 ---AAGTTCACACCAAGTCCGATGCTGGAGGTTTGGGGGCTGCTGGGAGAGTCTC 765
OY 418 PheSerTyrGly--ArgAlaProTyrProLysMetSerLeuLysGluValaSerGluAlaVala 437
DB 766 TTCTCATATGAGACCGGGCTCCGTACCTTAANAATGCTCTGAAGAAGGTGTGGAGGCCCT 825
OY 437 GlnLysGlyTyrArgMetGluProProGluGly--CysProGlyProValaHisValLe 456
DB 826 GGAAGAAGGGTACCGCATGGAACCCCGCGAGGCTGTGAGGCGCCCGCTGCAAGTCT 885
OY 456 UMet-SerSerCysTrpGluAlaGlu--ProAlaArgArgProProPhe 471
DB 886 CATGAACCACTGCTGGGAAGCAAAACCCGCGGCGCGGCGGCTTT 934
RESULT 6
AL580639/c 1035 bp mRNA linear EST 01-JUN-2003
LOCUS AL580639 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION AL580639 Homo sapiens cDNA clone CS0DJ007K05 3-PRIME, mRNA sequence.
ACCESSION AL580639
VERSION AL580639.2 GI:31318906
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1035)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 16, 2001 this sequence version replaced gi:12946853.
Contact: Genoscope
Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9238.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DJ007AF03NP1&cluster=9238.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DJ007AF03NP1.
 Location/Qualifiers
 1. 1035
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DJ007YK05"
 /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /cell_line="JURKAT"
 /clone_11b="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 166 a 327 c 304 g 213 t 25 others
 ORIGIN

Alignment Scores:
 Pred. No.: 4,72e-112 Length: 1035
 Score: 1330.00 Matches: 265
 Percent Similarity: 95.738 Conservative: 4
 Best Local Similarity: 94.318 Mismatches: 12
 Query Match: 49.798 Indels: 1
 DB: 9 Gaps: 0

us-09-977-260-2 (1-507) x ALS80639 (1-1035)

QY 227 GYTPLEuLeuAunLeuGlnHsLeuThrLeuGlyAlaGlnIleGlyGluGlyIuphe 246
 111
 1035 GCCTGCTMACTGAACTGACGATTTGACATGGGAGCAGATCGAGAGGAGAGTTT 976
 Db 247 GYAlaValLeuGlnGlyIupheGlyGlnIleGlyGluGlyIuphe 246
 111
 975 GGAAGCTCTCTGAGGAGGTGATGCTGGGCAAGAGTGGCTGAGAGATTCAGAGTGT 916
 QY 267 AspValThrAlaGlnAlaPheLeuAspGluThrAlaValMetThrIleMetGlnHsGlu 286
 111
 915 GATGTGACAGCCAGAGCTTCTCGAGAGGAGGCGGCTCATGACGAAGTGCACACGAG 856
 Db 287 AsnLeuValArgLeuGlnGlyValIleLeuHsGlnGlyLeuThrIleValMetGluHs 306
 111
 855 AACCTGCTGCTCTCTGAGGCTGATCGACACGAGGCTGATCAATTCATGAGGAC 796
 QY 307 ValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGlyValArgAlaLeuValAsnThr 326
 111
 795 GTGAGCAAGGCAACCTGCTGAATTTCTGCGGAGCCGGGCTGAGCCCTCGGAACACC 736
 Db 327 AlaGlnLeuLeuGlnPheSerLeuHsValAlaGluGlyMetGlyIupheGlnIuphe 346
 111
 735 GCTGAGCTCTGCTGAGTTTCTCTGACGCTGAGGCGGAGGAGGATGAGTACCTGGAGACAG 676
 QY 347 LysLeuValHisArgAspLeuAlaAlaGlnIleLeuValSerGlnAspLeuValAla 366
 111
 675 AAGCTTGTACACCGGACCTGCGCCGCCCAACATCTCTCTCAAGAACTGTGTGGCC 616
 Db 367 LysValSerAspPheGlyLeuAlaLysAlaGluArgGlyGlyLeuAspSerSerArgLeu 386
 111
 615 AAGNTCAGGCACTTTGGCTGGCCCAAGCCAGGAGGAGGAGTCAACGACCGGCTG 556
 QY 387 ProValLysTrpThrAlaProGluAlaLeuLysHsGlyLysPheThrSerLysSerAsp 406
 111
 555 CCGGTCAACTGGAGGCGCGGAGGCTCTCAACACGGAAGTTCACGCAAGTCGAT 496
 Db

QY 407 ValTrpSerPheGlyValLeuLeuTrpGluValAlaPheSerTrpGlyArgAlaProTrp 426
 111
 495 GTCTGAGATTTGGGTGGTCTGCTGAGAGTCTTCTCATATGAGAGGCTCCGTACCT 436
 Db 427 LysMetSerLeuGlnGlyValSerGluAlaValGluGlyLysGlyTrpArgMetGluProPro 446
 111
 435 AAATGTCACGTGAAGAGGTGTGGAGGCTTGAAGAGGAGTACCGCATGGAACCCGCC 376
 QY 447 GluGlyCysProGlyProValHisValLeuMetSerSerCysTrpGluAlaProAla 466
 111
 375 GAGGCGTTCACAGGCGCCGTCGACGTCATGACAGTGTGCTGGAGGAGGAGCCGCC 316
 Db 467 ArgArgProProPheArgLysLeuAlaGluLysLeuAlaArgGluLeuArgSerAlaGly 486
 111
 315 CCCCAGCCACCTTCACAAAGTGGCGGAGAACTGCGCCCGGGA-CTACGCAATGCAAGT 257
 QY 487 AlaProAlaSerValSerGlyGlnAlaPheAlaAspGlySerThrSerProArgSerGlnGlu 506
 111
 256 GCCCAGCCTCTCTCAAGGAGCAGAGCGGAGGAGGCTCACTGCGCCGCAAGCCAGAG 197
 Db 507 Pro 507
 111
 196 CCC 194

RESULT 7
 BX369553 986 bp mRNA linear EST 08-MAY-2003
 LOCUS BX369553 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
 DEFINITION Homo sapiens cDNA clone CS0DJ015Y001 5-PRIME, mRNA sequence.
 ACCESSION BX369553
 VERSION BX369553.1 GI:30457823
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9238.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0BAF0382F05.AF03591_2&cluster=9238.r.
 Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0BAF0382F05.AF03591_2.

FEATURES
 source 1. 986
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DJ015Y001"
 /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /cell_line="JURKAT"
 /clone_11b="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 207 a 270 c 321 g 187 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 2,93e-108 Length: 986
 Score: 1288.50 Matches: 288
 Percent Similarity: 88.25% Conservative: 5

Best Local Similarity: 86.75% Mismatches: 20
 Query Match: 48.24% Indels: 19
 DB: 13 Gaps: 3

US-09-977-260-2 (1-507) x BX369553 (1-986)

QY 195 HstYserLysAsp-LysGlyAlaIleCysThrLysLeuValArpProLysArgLysH 214
 DB 3 CATTACGCAAGGCAAGGCGCTATCTGCACCAACTGCTGAGCAACCAAGCGAACA 62
 QY 214 sglYthLysSerAlaGluGluGluLeuAlaArgAlaGlyTrpLeuLysLeuGlnH 234
 DB 63 CGGACCAAGTCCGCGCAGAGAGAGAGTGGCAGGCGGCTGTTACTGAACCTGCAGCA 122
 QY 234 sleuThrLeuGlyAlaGlnIleGlyGluGluGluPhgGlyAlaValLeuGlnGlyUty 254
 DB 123 TTTGACATTTGGGACACAGATCGAGAGGAGAGTGGAGCTTCTGCGAGGGTGA 182
 QY 254 rLeuGlyGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGlnAlaPhe 274
 DB 183 CTGGGGCAAAAGGTGGCCGCGAAGATATCAAGTGTGATGTGACACCCAGGCTTCT 242
 QY 274 uAspGluThrAlaValMetThrLysMetGlnHisGlnAsnLeuValArgLeuGlyUty 294
 DB 243 GAGAGAGAGCGCCCTCATGACGAGAGATGCAACACAGAACCTGCTCTCTGCGGCGT 302
 QY 294 lIleLeuHisGlnGlyLeuUtyrLysLeuMetGlnHisValSerLysGlnLysLeuValAs 314
 DB 303 GATCTTCACAGGAGGCTGTACATGTCATGAGACAGTGTGACAGGCAACCTGTGAA 362
 QY 314 nPheLeuArgThrArgGlyArgAlaValAsnThrAlaGlnLeuGlnPheSer 334
 DB 363 CTTTCTCGGAGCCCGGGGTGAGCCCTGCGTAACACCGCTAGTCTCTGCGGTTTCTCT 422
 QY 334 uHisValAlaGluGlyMetGlyUtyrLeuGlnSerLysLysLeuValHisArgAspLeu 354
 DB 423 GCAGGTGGCGAGGCGCATGTGAGTACCTGAGAGAGCAAGACCTGTGACACCGCAGCTGC 482
 QY 354 aAlaArgAsnIleLeuValSerGluAspLeuValAlaValSerAspPheGlyLeuVal 374
 DB 483 CGCCCGCAACATCTGTGCTCAGAGACCTGTGCGCAAGGTCCAGCATTTGGCTGCGC 542
 QY 374 alYsAlaGluArgLysGlyLeuAspSerSerArgLeuProValLysTrpThrAlaProG 394
 DB 543 CAAGCGGAGGAGGAGGCGTACACTCAAGCGGCTGCGCTCAAGTGGAGGCGCGCA 602
 QY 394 uAlaLeuLysHisGlyLysPheThrSerLysSerAspValTrpSer-PheGlyValLeu 414
 DB 603 GGCTCTCAACACCGGAGATTCAACACAGTGTGAGTGTGAGTGTGGGTGCTGC 662
 QY 414 eUTrpGluValPheSerTrp---GlyArgAlaProUtyrProLysMetSerLeuGlyUty 433
 DB 663 TCTGGGAGGCTTCTTCATGTGGAGGCGGCTCCGTCCCAAAATGTCATGAAAGG 722
 QY 433 alSerGluAla-ValGluLys-GlyTyrArgMetGluProProGluGly-CysProGlyP 452
 DB 723 TGTGGAGGCGCCGCGAGATAGGTTACCGCATGGAACCCCGAGGGCTGTGCAGGCC 782
 QY 452 roValHisValLeu-MetSerSer-CysTrpGluAlaGlu--ProAlaArgArgProPro 470
 DB 783 CCGGACACCTNCTTCATGAGAGCTTGTGGAGCGAAGCCCTGCGCGCGCCAGCC 842
 QY 471 PheArgLysLeu-AlaGluLysLeuAla-ArgGluLeu-ArgSerAlaGlyAlaProAla 489
 DB 843 TTTGGAACTGTTCGAGAACTTGTCCGAGAGCTTTCGCAATGGGCGGTGGCCAG 902
 QY 490 Ser-----ValSerGlyGlnAspAlaAspLys-----SerThrSerProArg 503
 DB 903 TCCCTTGGTTTAAGGGAGGAGGAGGAGGATGTTTCACTTTGTCTCCCGCAACCCAGG 962
 QY 504 SerGlnGluPro 507
 DB 963 AGTCCTGTCC 974

RESULT 8
 AL558929
 LOCUS
 DEFINITION
 ACCESSION
 AL558929
 VERSION
 AL558929.2
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1040)
 L.W.B., Gruber,C., Jessee,J. and Polayres,D.
 Full-length cDNA libraries and normalization
 Unpublished
 On Feb 15, 2001 this sequence version replaced gi:12903930.
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 9238.r For more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DJ007AF03QPI&cluster=9238.r. Contact :
 Feng Liang Email : fliang@life.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DJ007AF03QPI.
 Location/Qualifiers
 1..1040
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DJ007YK05"
 /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /cell_line="JURKAT"
 /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES

source

BASE COUNT 188 a 341 c 338 g 162 t 11 others
 ORIGIN

Alignment Scores:
 Pred. No.: 6.25e-103 Length: 1040
 Score: 1231.00 Matches: 231
 Percent Similarity: 98.72% Conservative: 1
 Best Local Similarity: 98.30% Mismatches: 3
 Query Match: 46.09% Indels: 1
 DB: 9 Gaps: 0

US-09-977-260-2 (1-507) x AL558929 (1-1040)

QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
 DB 335 ATGGCGGGGCGCAGGCTCTGTGGTTCTGGGCGGCATTTCACGGCTGTGATTCGCTGAG 394
 QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisTrpProProValSerAla 40
 DB 395 GAACCTCCCGGAGGAGCCCGCTTCTCCGAGCTGTGCAACCCCTCCGCTGTACGCC 454
 QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
 DB 455 AGGATGCCAACGAGGCGGTGGGCCCGGAGCAACCAAGTATCACCAATGCGAGCACACC 514
 QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValValThrIleLeuGluAla 80
 DB 515 CGCCCAAGCGAGGAGGAGTGGCTTCCGCAAGGCGCAGTGTGCACATCTCTGAGGCC 574

QY 81 CysGluAsnLysSerTrpTyrArgValLysHisThrSerGlyGlnGlyLeuLeu 100
|||||
Db 575 TGCAGAACAGAGCTGTATCCGCTCAGACACACCACTGACAGGAGCGCTG 634
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
|||||
Db 635 GCACCTGGGGGCTGGCGGAGCGGAGCCCTCTCCACAGACCCCAAGCTCAGCTCATG 694
QY 121 ProTrpPheHisGlyLysLysSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
|||||
Db 695 CCGGTGTTCCACGGGAGATCTCGGCGAGAGGCTCTCCAGCAGCTGACGCTCCGAG 754
QY 141 AspGlyLeuPheLeuValArgGlnSerAlaArgHisProGlyAspTyrValLeuCysVal 160
|||||
Db 755 GATGGCTGTTCCTGGTCCGAGGAGTCCGCGCCACCCCGGAGCTACGCTCTGTGCTG 814
QY 161 SerPheGlyArgAspValLleHisTyrArgValLeuHisArgAspGlyHisLeuThrLle 180
:::
Db 815 ACCGT-GGGCGGACGCTCATCTCAGCTACGCGTGTGTCACGCGGACGCTCAGATC 873
QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLysAspLys 200
|||||
Db 874 GATGAGCGCTGTTCTTCTCCTCAACCTCATGACATGTGTGACATTCACAGACAGACAG 933
QY 201 GlyAlaLleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
|||||
Db 934 GCGGCTATCTGCACCAAGCTGTGTAGACCAAGCGAAGACGAGGACCAAGTGGCGGAG 993
QY 221 GluGlnLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeu 235
|||||
Db 994 GAGGAGCTGGCGAGGCGGCTGTACTGACCTGAGCACTTGG 1038

RESULT 9
BM450510 1129 bp mRNA . linear EST 05-FEB-2002
DEFINITION
AGENCOURT_6394260 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5494913
5', mRNA sequence.
BM450510
BM450510.1 GI:18499550
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1129)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12120 row: p column: 18
High quality sequence stop: 600.
Location/Qualifiers
1. 1129

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5494913"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-Sport6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 233 a 345 c 369 g 182 t

ORIGIN

Alignment Scores:
Pred. No.: 1,82e-102 Length: 1129
Score: 1226.50 Matches: 271
Percent Similarity: 87.46% Conservative: 8
Best Local Similarity: 84.95% Mismatches: 26
Query Match: 45.92% Indels: 14
DB: 12 Gaps: 8
US-09-977-260-2 (1-507) x BM450510 (1-1129)

QY 89 ValLysHisThrSerGlyGlnGlyLeuLeuAlaAlaGluArgGluArg 108
|||||
Db 6 GTCAAGCACACACACCACTGACAGAGGAGGCTGTGACAGTGGCGGCTGGAGCGG 65
QY 109 GluAlaLeuSerAlaAspProLysLeuSerLeuMetProTrpPheHisGlyLysLleSer 128
|||||
Db 66 GAGGCCCTCTCCGACAGACCCCAAGCTCAGCTCATGCGGTGTTCCACGGGAGATCTCG 125
QY 129 GlyGlnGluAlaValGlnGlnLeuGlnProProGluAspGlyLeuPheLeuValArgGlu 148
|||||
Db 126 GCCCAGAGGCTGTCCAGACGCTGACACCTCCGAGATGGCTGCTCCGTGCGGAG 165
QY 149 SerAlaArgHisProGlyAspTyrValLeuCysValSerPheGlyArgAspValLleHis 168
|||||
Db 186 TCCGCGGCGCACCCCGGCTACGCTGTGCTGTGAGCTTTGGCCGACGCTCATCCAC 245
QY 169 TyrArgValLeuHisArgAspGlyHisLeuThrLleAspGluAlaValPhePheCysAsn 188
|||||
Db 246 TACCGCTGTGACCGCGGCGGCTCAGCTCAGATCATGATGAGCGGTCTTCTGACAC 305
QY 189 LeuMetAspMetValGlnHisTyrSerLysAspLysGlyAlaLleCysThrLysLeuVal 208
|||||
Db 306 CTCATGACATGTGTGAGCATTTAGACATTTGGAGCACAGATCGGAGAGGAGGTTGGAGCT 365
QY 209 ArgProLysArgLysHisGlyThrLysSerAlaGlnGluLeuAlaArgAlaGlyTrp 228
|||||
Db 366 AGACCAAGGCGAAGACCGGACCAAGTCGCGGAGAGAGCTGCGCACGCGGCTGG 425
QY 229 LeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnLleGlyGlnGlyPheGlyAla 248
|||||
Db 426 TTACTGACCTGACGACATTTGACATTTGGAGCACAGATCGGAGAGGAGGTTGGAGCT 485
QY 249 ValLeuGlnGlyLysTrpLeuGlnGlyLysValAlaValLysAsnLleLysCysAspVal 268
|||||
Db 486 GTCTCTCAGGCTGAGTACCTGGGCGAAGAGTGGCGTGAAGATTCAGATGTGATG 545
QY 269 ThrAlaGlnAlaPheLeuAspGlyThrAlaValMetThrLysMetGlnHisGluAsnLeu 288
|||||
Db 546 ACAGCCCGAGGCTTCTGCGAGCGAGCGGCTGATGACGAAGTGCACACGAGAACCTG 605
QY 289 ValArgLeuLeuGlyAlaLleLeuHisGln-GlyLeuTyrLleValMetGlnHisValSe 308
|||||
Db 606 GTCCGCTCTCTGGCGCTGATCTCGACACAGGAGGCTGTACATTTGTATGAGCACCTGAG 665
QY 308 rLysGlyAsnLeuValAsnPheLeuArg-ThrArgGlyArgAla-LeuValAsnThrAla 327
|||||
Db 666 CAAGGCGAACCTGTGAGCTTCTGCGGAGACCGGCGGTGAGGCGCTCGGAACACGCTG 725
QY 328 Gln-LeuLeuGln---PheSerLeuHisValAlaGlnGly---MetGlnTyrLeu---G1 344
|||||
Db 726 CAGCCTCTCGAGGATTTCTGTGCCCCGCGGCGGACATTGGAGTTTACCTGCGGAGA 785
QY 344 uSerLysLysLeuValHisArg-AspLeuAla-----AlaArgAsnLleLeuValSerG 362
|||||
Db 786 ACAAGAAACCTTTGTACACCGGAGCTGCCCCCGGCAATTCCTGGCCCCCAAG 845
QY 362 LuAspLeuValAlaLysValSer---AspPheGlyLeuAlaLysAlaGluArg---LysG 380
|||||
Db 846 GACCTGGGGGCGCAAGTGCAGCGGACCTTGGGCGCTGCGCAAAAGCCCAAGCGAGG 905
QY 380 LysLeuAsp---SerSerArgLeu---ProValLysTrpThrAlaPro 393

Db 906 GGCTAAATTCAGCCCGCTGCGCTCAGAGGGGAGCGGCC 952
||||| :|||
RESULT 10
LOCUS AL558805 1032 bp mRNA linear EST 31-MAY-2003
DEFINITION Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
ACCESSION AL558805 Homo sapiens cDNA CS0DJ015X001 5-PRIME, mRNA sequence.
VERSION AL558805.2 GI:31280603
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1032)
AUTHORS L.M.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12903685.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9238.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ015AH01QPL&cluster=9238.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradey Avenue Genoscope sequence ID : CS0DJ015AH01QPL.
Location/Qualifiers
1..1032
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015X001"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 187 a 338 c 334 g 164 t 9 others
ORIGIN
Alignment Scores:
Pred. No.: 1.49e-101 Length: 1032
Score: 1216.00 Matches: 227
Percent Similarity: 98.70% Conservative: 1
Best Local Similarity: 98.27% Mismatches: 3
Query Match: 45.53% Indels: 1
Gaps: 0
US-09-977-260-2 (1-507) x AL558805 (1-1032)
QY 1 Metalaglyargylserleuvalsertrpargalaphenisglycysaspseralaglu 20
Db 338 ATGGCGGGGAGGAGCTCTGCTGTTCTGGGGGCGCATTCACGGCTGTGATTCGCGAG 397
QY 21 Gluleuproargvaliserproargpheleuargalatrphispropiovalserala 40
Db 398 GAACCTTCCCGGGGAGCGCCCTCTCGACCTGGCGACCCCTCCCGCTCGAGCC 457
QY 41 ArgmetprohrargatrgtrpalaaproglythrincysilethrlyscysgluHsthr 60
Db 458 AGGATGCCACGAGGCGCTGGGGCGCCGCGCACCAAGTATACCAATGGCAGACACC 517
QY 61 Argprolysproglygluleuvalaphenarglysglyspalvalthrileugluala 80
|||||

Db 518 CGCCCAAGCCAGGAGAGCTGGCTCCGCAAGGCGAGCTGTACCATCTCGAGGCC 577
QY 81 Cysglunsluyssetrpttrpargvallyshsthrserglyngluuglyleu 100
Db 578 TGCAGAGAACAGACTGTGTACCGGTCAAGACACACCAATGAGAGAGGCTGTG 637
QY 101 AlaalaglyalalauarggluarglualaleuSERalaappprolysluSERleuMET 120
Db 638 GCAGCTGGGGCGCTCGCGGAGCGGAGCCCTCTCCGACAGCCCAAGCTCAGCTCATG 697
QY 121 ProtrpPheHISglylysIleSerGlyngluvalaValngluInleuGlnProProglu 140
Db 698 CCGTGTTCACAGGAGAGATCTCGGCGAGAGGCTGTCCAGAGCTCAGCTCAGCTCCGAG 757
QY 141 AspGlyleupheleuvalarggluSERalarghstprsglyasptlyvalleucysval 160
Db 758 GATGGCGCTTCTGTGTCGGGAGCTCGCGGCCACCCCGGAGCTAGCTGTGCTGCTG 817
QY 161 SerPheglYargaspvalIleHIStrYargvalleuHISArgspgLYHISleuThrIle 180
Db 818 ASTTTGGCGCGAGCTATCTACTACCGCTGTGCGACCGGAGGAGGAGCTCAGATC 877
QY 181 AspGluAlaValPhePheCysAsnleuMetaspMetValGluHISrYserLYsAspLYs 200
Db 878 GATGAGGCGCGTCTTCTGTCAACCTCATGACATGTGGAGCATTTACAGCAGACAG 937
QY 201 GlyAlaIleCysThrlysluvalargprolysarglyshsglythrlyseralaglu 220
Db 938 GGCCTATCTTCACCAACCTGTGAGACAAAGCGGAACACGGAGACCAAGTCGGCCAG 997
QY 221 GluGluLeuAlaArgAlaGlyTrpLeuAsn 231
Db 998 GA-GAGCTGGCCAGAGCGGCTGTACTGAAAC 1029
RESULT 11
Bg744770 906 bp mRNA linear EST 15-MAY-2001
LOCUS Bg744770/c
DEFINITION 602722891J1 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:4849175 3',
mRNA sequence.
ACCESSION Bg744770
VERSION Bg744770.1 GI:14055423
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 906)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMU at:
http://image.llnl.gov
Plate: L16M168 row: f column: 24
High quality sequence start: 21
High quality sequence stop: 845.
Location/Qualifiers
1..906
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4849175"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_106"
/note="Organ: blood; Vector: pOT87; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC library."

BASE COUNT 142 a 306 c 281 g 176 t 1 others
ORIGIN

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
5, 47e-101	906	1209.00	241	1	5	1	0
Percent Similarity:		97.98%					
Best Local Similarity:		97.57%					
Query Match:		45.26%					

US-09-977-260-2 (1-507) x BG744770 (1-906)

```

OY 261 VALLYSANLLELYSCYSPVALTHRALAGLALPHELEUNSPGLTHRALAVALMET 280
    |||||:::|||||
DB 899 GTMAAGAAATATCAAGGGTGTGG-ACAGCCAGGCGTTCCTGGAGCAGACGCCGTCATG 841

OY 281 ThrlyMetGlnHisGlnLeuValArgLeuGlyValIleuHisGlnGlyLeu 300
    |||||
DB 840 ACAGAGATGCACAGACAGAACCTGGTGGCTCTCGGCGTGATCTGCACCGAGGGCTG 781

OY 301 TyrIleValMetGlnHisValSerIleGlyAsnLeuValAsnPheluarqThrArgly 320
    |||||
DB 780 TACATTGTCTAGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 721

OY 321 ArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGlnGlyMet 340
    |||||
DB 720 CGAGCCCTGGTGAACCGCTCAGCTCGACGATTCTCTGACACCTGGCCGAGGCGCATG 661

OY 341 GlnTyrLeuGlnSerIleValLeuValHisArgAspLeuAlaIleArgAsnIleLeuVal 360
    |||||
DB 660 GAGTACCTGGAGAGCAAGACCTGTGTGACCGGACCTGGCCGCCGCAACATCTGTGTC 601

OY 361 SerGluAspLeuValAlaValSerAspPheGlyLeuAlaValSerIleGlyGly 380
    |||||
DB 600 TCAGAGACCTGGTGGCCAGGTCAGCGACGATTGGCGTGGCCAAACCGGAGGAGGG 541

OY 381 LeuAspSerSerArgLeuProValIleThrAlaProGlnAlaLeuValHisGlyIle 400
    |||||
DB 540 CTAGACTCAAGCGGCTGCCCGCTCAAGTGAACGCGCCGAGGCTCTCAAAACGCGGAG 481

OY 401 PheThrSerIleSerAspValTrpSerPheGlyValIleLeuThrProGlnAlaPheSer 420
    |||||
DB 480 TTCACACGACAGTGGATGTGTGGAGTTTGGGGTCTGTCTGGAGAGGTCTTCTCATAT 421

OY 421 GlnArgAlaProTyrProIleMetSerLeuValSerIleGlnValAlaValGlnGly 440
    |||||
DB 420 GGACGGGCTCCGACCTAAATGTCACGAAAGAGGTGGAGGCGCGGAGGAGGG 361

OY 441 TyrArgMetGlnProProGlnGlyCysProGlyProValHisValLeuMetSerSer 460
    |||||
DB 360 TACCGATGGAAACCCCGCAGGGCTGTCCAGGCGCGCTGACCTCTCATAGACACTGC 301

OY 461 TrpGlnAlaGlnProAlaArgProProPheArgIleLeuAlaGlnIleValSer 480
    |||||
DB 300 TGGGAGGCGAGAGCCCGCGCGCGGCGACCTTCGCAAACTGGCCAGAAAGCTGGCGCG 241

OY 481 GlnLeuArgSerAlaGlnAlaProAlaSerValSerGlyIleAspAlaSpGlySerThr 500
    |||||
DB 240 GAGCTACGACAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 181

OY 501 SerProArgSerGlnGlnPro 507
    |||||
DB 180 TCGCCCGGAAAGCCAGAGCC 160
  
```

RESULT 12
AL580543/c

LOCUS AL580543 1012 bp mRNA linear EST 01-JUN-2003
DEFINITION Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
ACCESSION AL580543
VERSION AL580543.2 GI:31318812
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 16, 2001 this sequence version replaced g1:1294667.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

FEATURES
source
1..1012
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015Y001"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
Faraday Avenue Genoscope sequence ID: CS0DJ015AH01NP1.
Location/Qualifiers
1..1012

BASE COUNT 158 a 346 c 277 g 194 t 37 others
ORIGIN

Alignment Scores:
Pred. No.: 9, 77e-101 Length: 1012
Score: 1207.00 Matches: 253
Percent Similarity: 93.77% Conservative: 3
Best Local Similarity: 92.67% Mismatches: 15
Query Match: 45.19% Indels: 6
DB: 9 Gaps: 1

US-09-977-260-2 (1-507) x AL580543 (1-1012)

```

OY 235 LeuThrLeuGlnValAlaGlnIleGlyGlnGlyPheGlyAlaValLeuGlnGlyTyr 254
    :::::|||||
DB 1006 ATTACATTGGAGSACAGA---TCGAGAGAGGAGATTGGACT-CTCTGCA-GGTGAGTAC 952

OY 255 LeuGlnGlnValAlaValIleValSerAsnIleValSerAspValThrAlaGlnAlaPheLeu 274
    |||||
DB 951 CT-GGGCAAAAGGTGGCCCGTGAAGAAATATCAATGTGATGTGCACACCCAGGCTTCTCTG 893

OY 275 AspGlnThrAlaValMetThrIleValMetGlnHisGlnAsnLeuValArgLeuGlnGlyVal 294
    |||||
DB 892 GACGACAGCGCGCTATGAGAGAGCAAGCAACGAAACCTGGTGGTGGTGGCGGCGTG 833

OY 295 IleLeuHisGlnGlyLeuTyrIleValMetGlnHisValSerIleGlyAsnLeuValAsn 314
    |||||
DB 832 ATCTCTCACACAGGGGCTGTATCATTCATGAGACAGTGAAGAGGCAACCGTGTGAAC 773

OY 315 PheLeuArgThrArgGlnArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeu 334
    |||||
DB 772 TTTCTGCGAGCCCGGGGTGGAGCCCTCGTGAACACCGCTCAGCTCTGAGATTCTCTG 713
  
```



```

Db      757 GCAGGGCGGTCCAGCTCTGTCAAGCAGA 786
|||||
RESULT 14
LOCUS   AL567349 949 bp mRNA linear EST 31-MAY-2003
DEFINITION AL567349 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
ACCESSION CS00D027YD13 3-PRIME, mRNA sequence.
VERSION   AL567349
KEYWORDS  AL567349 GI:31290254
SOURCE    EST.
ORGANISM  Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS  Li, W. B., Gruber, C., Jesse, J., and Polayes, D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished
COMMENT   On Feb 16, 2001 this sequence version replaced gi:12920618.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9238.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DF027CB07NP1&cluster=9238.r. Contact:
Peng Liang Email: liang@lifetech.com URL:
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS0DF027CB07NP1.
Location/Qualifiers
1. 949
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF027YD13"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: PCWSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime cDNA
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the PCWSPORT 6
vector. Library was not normalized."
BASE COUNT 144 a 279 c 290 g 212 t 24 others
ORIGIN
Alignment Scores:
Pred. No.: 4.45e-96 Length: 949
Score: 1156.00 Matches: 229
Percent Similarity: 94.90% Conservative: 13
Best Local Similarity: 89.80% Mismatches: 13
Query Match: 43.28% Indels: 1
DB: 9 Gaps: 0
US-09-977-260-2 (1-507) x AL567349 (1-949)
QY 253 GUTYTLGGLGSLNLYSVALAVALYASNLGLYSCYSAPVALTHRALAGNALA 272
|||||
Db 948 GAGTACCTGGGGGAAAAGGTGGCCGTGAGAAATATCAAGTGTGTGACAGCCCGAGGCC 889
QY 273 Pheleuaspeluthralavalmetthrlysmetglnhsglulansleuvalargleu 292
|||||
Db 888 TTCCTGAGACGAGCGCCGTGATGAGAGATGCAACAGAGAACTGGTCTCTCC 829
QY 293 GLYVALILEUHLISGLNGLYLEUTRYLLEVALMETGLUHSVALSERLYSGLYASNL 312
|||||
Db 828 GGGGTATCTCTGACACGAGGGCGTGTACATGTGTATGAGAGACCTGAGCAAGGCAACCTG 769
QY 313 Valaspheleuargthrargglarvalaleuvalasnthralaglnleuaglnpe 332
|||||

```

```

Db      768 GTGAACTTCTGCGGACCCGGGGTCGAGCCCTGCTGAACACCCCTCAGCTCTGCACTTT 709
QY 333 Serleuhsvalalaglglumetglutrylleugluserlyslsleuvalhlsargasp 352
|||||
Db 708 TCTGTGACGTGGCCGAGGGCGATGAGTACTGTGAGAGCAAGAAAGCTGTGTGACCGCGAC 649
QY 353 LeuAlAlAArgasnilleuvalsergluaspleuvalalalysvalseraspphegly 372
|||||
Db 648 CTGGCCCGCCGCAACATCTGATCTCAGAAACCTGTGGCCGCAAGCAACACTTTGGC 589
QY 373 LeuAlAlAArgasnilleuvalsergluaspleuvalalalysvalseraspphegly 392
|||||
Db 588 GTGGCAAAWCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
QY 393 ProgluAlAlAArgasnilleuvalsergluaspleuvalalalysvalseraspphegly 412
|||||
Db 528 CCGAAGCTCTCAACACGAGGAAATTCACCAACAATCATATCTGATTTTGGGGTG 469
QY 413 LeuLeuTrpGluValPheSerTyrglyArgAlaProTyrProLysMetSerLeuLysGlu 432
|||||
Db 468 CTGCTCTGGGARATCTTCTCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAR 409
QY 433 ValSerGluAlAlAArgasnilleuvalsergluaspleuvalalalysvalseraspphegly 452
|||||
Db 408 GTGTGAGACCGGTGAGAGARRRGTACCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 349
QY 453 ValHisValLeuMetSerSerCysTrpGluAlAlAArgProAlaArgArgProphearg 472
|||||
Db 348 GTGACAGTCTCTCATGAGACACCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 289
QY 473 LysLeuAlAlAArgasnilleuvalsergluaspleuvalalalysvalseraspphegly 492
|||||
Db 288 AACTGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 230
QY 493 GLYGLASPALASPGLYSERTHRNSERPROARGSERGLNLU 507
Db 229 GRCGAGACCGCCGACRGTCTCACCTGCGCCGAGAGAGAGAGAGAGAGAGAGAGAG 185
RESULT 15
LOCUS   BI912704 837 bp mRNA linear EST 16-OCT-2001
DEFINITION 60317635F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5240700 5',
mRNA sequence.
ACCESSION BI912704
VERSION   BI912704.1 GI:16176923
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS  NIH-MGC http://mgs.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
Email: cgaps-f@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L16M1606 row: P column: 13
High quality sequence stop: 770.
Location/Qualifiers
1. 837
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5240700"
/lab_host="DH10B"
/clone_lib="NIH_MGC_121"
FEATURES
Source

```

/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH-MGC Library."

BASE COUNT 157 a 267 c 291 g 122 t
ORIGIN

Alignment Scores:
Pred. No.: 1.37e-94 Length: 837
Score: 1139.00 Matches: 230
Percent Similarity: 90.31% Conservative: 3
Best Local Similarity: 89.15% Mismatches: 12
Query Match: 42.64% Indels: 16
DB: 12 Gaps: 1

US-09-977-260-2 (1-507) x B1912704 (1-837)

```
OY 2 AlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 21
DB 88 GCGGGGCGCGGC-----TCGGGGCGCGCC 111
OY 22 LeuProArgValSerProArgPheLeuArgAlaTrpHisProProProValSerAlaArg 41
DB 112 CCGAGCAGGTGAGCCCTCTCCGAGCCTGGACCCCTCCGCTCAAGCCAGG 171
OY 42 MetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGlnHisThrArg 61
DB 172 ATGCCAAGAGGCGGTGGCCCGCGGACCCAGCTATCACCMAATGGAGACACCCGC 231
OY 62 ProlYsProGlyGluLeuAlaPheArgLysGlyAspValValThrIleLeuGlnAlaCys 81
DB 232 CCCAAGCCAGGGAG-CTGGCTTCGCGAAGGGCGACGTGCACCATCTCGAGGCTGC 290
OY 82 GluAsnLysSerTrpTyrArgValLysHisThrSerGlyGlnGlyLeuLeuAla 101
DB 291 GAGAACAGAGCTGTGTACCGCTCAGACACCAAGTGAAGAGGGCTGTGCA 350
OY 102 AlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProlYsLeuSerLeuMetPro 121
DB 351 GCTGGGGCGCTGCGGAGCGGAGGCCCTCTCCGAGACCCCAAGCTCAGCTCATGCCG 410
OY 122 TrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGluAsp 141
DB 411 TGGTTCCACGGGAGATCTCGGCGCAGAGGCTGCCAGAGCTGCAGCTCCGAGGAT 470
OY 142 GlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuCysValSer 161
DB 471 GGGCTGTCCTGTGGGAGATCTCGGCGCAGAGGCTGCCAGCTCTGTGCGTGAAC 530
OY 162 PheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIleAsp 181
DB 531 TTGGCGCGGAGCATCATCTACCGCTGCTGCACCGCAGCG-CACTTCACATCGAT 589
OY 182 GluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLysAspLysGly 201
DB 590 GAGGCGGTCTCTCTGCACCTCATGTGAGCATGTGAGCATTCACAGCAAGACAAATGCC 649
OY 202 AlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGluGlu 221
DB 650 GCTATCTGCACCAAGCTGTGAGCAACCAAGGAAACAGGAGCAACGTGCGCGAGGAG 709
OY 222 GluLeuAlaArgAla-GlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnI 241
DB 710 GAGCTGGCGCAGGGGGGCTGTACTGACCTGAGCATTTGACATTGGAGACACAGAT 769
OY 241 eGlyGluGlyLubPheGlyAlaValLeuGlnGlyGluTyrLeuGlyGlnLys 258
|||||
```

DB 770 CGAGAGAGAGAGAGTTGGAGCTGC-CTGCAGGCTGAGTACTGGGCAAAAGG 820
Search completed: August 1, 2003, 22:37:07
Job time : 3256 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 2, 2003, 04:55:27 : Search time 971 Seconds
(without alignments)
1077.180 Million cell updates/sec

Title: US-09-977-260-2
2671
Perfect score: 1 MAGRSLVSWRAFGHDSAE.....PASVSGDADGSTSPRSGEP 507
Sequence:

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1439767 seqs, 1031500376 residues
Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-O=/cgn2_1/USPTO/US09977260/runat_29072003_092742_14668/app.query.fasta_1.647
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=tnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=DIsum62
-TRANS-human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pcp -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09977260.ecgn.1.1.333 @runat_29072003_092742_14668
-NCPU=6 -ICPU=3 -NO_MAP -LARGEOUT -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match Length	ID	Description
No.					
1	2671	100.0	2000	US-09-977-269-1	Sequence 1, Appl1

2	2671	100.0	2000	US-09-977-260-1	Sequence 1, Appl1
3	2671	100.0	2000	US-09-977-261-1	Sequence 1, Appl1
4	2671	98.0	1989	US-10-084-817-341	Sequence 341, App
5	1247.5	46.7	2187	US-09-954-531-188	Sequence 188, App
6	1247.5	46.7	2187	US-10-298-377A-1	Sequence 1, Appl1
7	1247.5	46.7	2420	US-10-177-593-87	Sequence 87, Appl1
8	725.5	27.2	2015	US-09-954-456-1983	Sequence 1983, App
9	725.5	27.2	2015	US-10-007-010-3	Sequence 3, Appl1
10	715.5	26.8	1911	US-09-917-800A-1611	Sequence 1611, App
11	710	26.6	4517	US-10-298-377A-3	Sequence 3, Appl1
12	709.5	26.6	2298	US-10-175-523-50	Sequence 50, Appl1
13	701.5	26.3	5527	US-09-880-107-3710	Sequence 3710, App
14	681.5	25.5	2354	US-09-967-768A-300	Sequence 300, App
15	654.5	24.5	2770	US-09-977-269-5	Sequence 5, Appl1
16	654.5	24.5	2770	US-09-977-260-5	Sequence 5, Appl1
17	654.5	24.5	2770	US-09-977-261-5	Sequence 5, Appl1
18	654.5	24.5	2663	US-09-954-456-1631	Sequence 1631, App
19	654.5	24.5	7607	US-09-982-610-19	Sequence 19, Appl1
20	633.5	23.7	4221	US-10-220-801-8	Sequence 8, Appl1
21	630.5	23.6	6383	US-09-954-531-405	Sequence 405, App
22	614.5	23.0	2560	US-10-045-302-1	Sequence 1, Appl1
23	608.5	22.8	2582	US-10-220-801-4	Sequence 4, Appl1
24	608.5	22.8	2674	US-10-003-295-1	Sequence 1, Appl1
25	607.5	22.7	2451	US-09-771-161A-4	Sequence 4, Appl1
26	607	22.7	2468	US-10-045-202-3	Sequence 3, Appl1
27	587	22.0	2449	US-10-220-801-11	Sequence 11, Appl1
28	587	22.0	2456	US-10-186-399-1	Sequence 1, Appl1
29	587	22.0	2500	US-09-977-269-3	Sequence 3, Appl1
30	587	22.0	2500	US-09-977-260-3	Sequence 3, Appl1
31	587	22.0	2500	US-09-977-261-3	Sequence 3, Appl1
32	586.5	22.0	2520	US-09-953-047-3	Sequence 3, Appl1
33	582.5	21.8	3829	US-09-953-047-10	Sequence 10, Appl1
34	579.5	21.7	2604	US-10-171-581-317	Sequence 317, App
35	578.5	21.7	2627	US-09-962-436-313	Sequence 313, App
36	576.5	21.6	2184	US-09-728-952-82	Sequence 82, App
37	576.5	21.6	3454	US-10-044-090-48	Sequence 48, App
38	576.5	21.6	3726	US-09-925-302-271	Sequence 271, App
39	563	21.1	3593	US-10-220-801-6	Sequence 6, Appl1
40	563	21.1	3663	US-09-919-172-84	Sequence 84, Appl1
41	558.5	20.9	3914	US-10-044-090-148	Sequence 148, App
42	556	20.8	2574	US-09-735-103-2	Sequence 2, Appl1
43	556	20.8	2574	US-10-045-428A-2	Sequence 2, Appl1
44	555	20.8	3306	US-09-954-556-10	Sequence 10, Appl1
45	551.5	20.6	2433	US-10-037-270-830	Sequence 830, App

ALIGNMENTS

RESULT 1
US-09-977-269-1
Sequence 1, Application US/09977269
Patent No. US20020082037A1
GENERAL INFORMATION:
APPLICANT: DILLRICH, AXEL
APPLICANT: GISHITZKY, MIKHAIL
APPLICANT: SURES, IRMINGARD
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
FILE REFERENCE: 038602/1260
CURRENT APPLICATION NUMBER: US/09/977,269
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 08/232,545
PRIOR FILING DATE: 1994-04-22
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 2000
TYPE: DNA
ORGANISM: Unknown Organism
FEATURE:
NAME/KEY: CDS
LOCATION: (258)..(1778)
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte

OTHER INFORMATION: kinase 1
US-09-977-269-1

Alignment Scores:

Pred. No.:	1,05e-266	Length:	2000
Score:	2671.00	Matches:	507
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-977-260-2 (1-507) x US-09-977-269-1 (1-2000)

```

QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
DB 258 ATGCGGGGGGAGGAGCTCTGTTCTCGCGGCAATTCACGGCTGTGTTCTGCTGAG 317
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProAlaSerAla 40
DB 318 GAACCTTCCCGGGGTGAGCCCCCTTCTCCGAGCCTGCGACCCCTCCCTCTCAGCC 377
QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysTleThrIleCysGluHisThr 60
DB 378 AGGATGCCAAGAGGGGCTGGCCCGGCGACCCAGTGTATCCAAATCCGAGCACACC 437
QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaThrIleLeuGluAla 80
DB 438 CGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGGGAGCTGGTCCACCATCTCGAGGCC 497
QY 81 CysGluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGlyLeuLeu 100
DB 498 TGGGAGAACAGAGCTGTACCGCGTCAACACCAACAGTGTGAGAGAGGGGCTGCTG 557
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
DB 558 GCGAGTGGGGGCTGGGGGAGCGGAGGCCCTCTCCGCAAGCCCAAGCTCAGCTCATG 617
QY 121 ProTrpPheHisGlyLysLysSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
DB 618 CCTGTGTTCCAGCGGAGATCTGGGCGCAGGAGGCTGTCCAGCAGCTGACGCTCCGAG 677
QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysVal 160
DB 678 GATGGGCTGTCTGTGGGGGAGTCCGCGCCGCCACCCCGGCACTACGCTGCTGCTG 737
QY 161 SerPheGlyArgAspValIleHisTrpArgValIleHisArgAspGlyHisLeuThrIle 180
DB 738 AGCTTTGGCGCGAGCATCCACTACCGGTGCTGACCGCGAGCCGACCTCAGCATTC 797
QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTrpSerLysAspLys 200
DB 798 GATGAGCGCGTGTCTTCTGCAACCTCATGACATGATGAGCATTAACAGCAAGCAAG 857
QY 201 GlyAlaIleCysThrIleValLeuArgProLysArgLysHisGlyHisThrIleSerAlaGlu 220
DB 858 GGGCGTATCTGCAACAGCTGTGAGACCAAGCGGAAACCGGAGCAAGTGGCGGAG 917
QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrIleGlyAlaGln 240
DB 918 GAGGAGCTGGCCAGGGGCGGCTGTACTGAACTGACGATTTGACATTGGGAGCAGAG 977
QY 241 IleGlyGluGlyLeuPheGlyAlaValLeuGlnGlyLutTrpLeuGlyLysValAla 260
DB 978 ATCGGAGAGGAGAGTGTGAGCTGTCTGCAAGGAGTGAAGTCTGGGGGCAAGGTGGCC 1037
QY 261 ValIleAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
DB 1038 GTGAGAGATATTCAGAGTGTGAGACAGCCAGGCTTCTCTGACAGAGAGCGCGTCAAG 1097
QY 281 ThrLysMetGlnHisGlnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
DB 1098 ACGAAGATGCAACAGCAAGAACCTGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1157

```

```

QY 301 TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgTrpArgLys 320
DB 1158 TACATGTGATGAGAGCAGTGTGAGCAAGGGCAACCTGGTGAACCTTCTGCGGACCGGCGGT 1217
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGluGlyMet 340
DB 1218 CGAGCCCTCTGTGAACCGCTGACCTGTGCAAGTCTTCTGTGACGTGGCGGAGGCATG 1277
QY 341 GluTrpLeuGlnSerLysLysLeuValHisArgAspLeuAlaIleArgAsnIleLeuVal 360
DB 1278 GAGTACCTGGAGAGCAAGAAAGCTTGTGACCCGCGGACCTGGCCCGGCAACATCTGTGTC 1337
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
DB 1338 TCAGAGAGCTGTGGGCAAGGTGAGAGATTTGGCTGTGCGCAAGCGGAGCGGAGGGG 1397
QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
DB 1398 CTGAGCTCAAGCCGCTGCGCGTCAAGTGAAGCGGCGCCGAGGCTCTCAAAACAGGGAG 1457
QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyr 420
DB 1458 TTCACCAAGCAAGTCGATGTCTGAGTTTGGGGTCTGCTGTGGAGAGTCTTCTCATAT 1517
QY 421 GlyArgAlaProTrpProLysMetSerLeuLysGluValSerGluAlaValAlaGlyLysGly 440
DB 1518 GAGAGGCTTCCGCTTAAATGCTCACTAAAGAGGTGTGAGAGCGCGTGGAGAGGGG 1577
QY 441 TyrArgMetGlnProProGlnGlyCysProGlyProValHisValLeuMetSerCys 460
DB 1578 TACCGATGGAACCCCGGAGGCTGTCCAGGCGCGGTGACGTCTCTCATGACACCTGTG 1637
QY 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
DB 1638 TGGGAGGAGAGGCGCGCGCGGCGGCACTTCCGCAACTGCGGAGAAAGTGGCGCGG 1697
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
DB 1698 GAGCTACGCAAGTGTGAGTGTGCGGCTGCTGTCTGAGGAGGAGAGCGCGGCTCCAGC 1757
QY 501 SerProArgSerGlnGluPro 507
DB 1758 TCGCCCCGAAAGCCAGGAGGCC 1778

RESULT 2
US-09-977-260-1
: Sequence 1, Application US/09977260
: Publication No. US20020192790A1
: GENERAL INFORMATION:
: APPLICANT: ULIRICH, AXEL
: APPLICANT: GISHIZKY, MIKHAEL
: APPLICANT: SURES, IRMINGARD
: TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
: FILE REFERENCE: 038602/1260
: CURRENT APPLICATION NUMBER: US/09/977, 260
: PRIOR FILING DATE: 2001-10-16
: PRIOR APPLICATION NUMBER: 08/232, 545
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 2000
: TYPE: DNA
: ORGANISM: Unknown Organism
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (258)..(1778)
: FEATURE:
: OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
: OTHER INFORMATION: kinase 1
US-09-977-260-1
Alignment Scores:

```

Pred. No.: 1.05e-266 Length: 2000
 Score: 2671.00 Matches: 507
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 10

US-09-977-260-2 (1-507) x US-09-977-260-1 (1-2000)

Qy 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
 Db ATGGGGGGGAGGCTCTGGTTTCCTGGGGGCGATTTCACGGCTGATTCGTGAG 317
 Qy 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
 Db GAATCTCCCGGGGAGAGCCCGCTCTCCGAGCCTGGACCCCGCTCCCGCTCAGCC 377
 Qy 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
 Db AGGATGCCAAGAGGCGCTGGGCGCGGACCCAGTGTATCACCMAATGGAGCACACC 437
 Qy 61 ArgProLysProGlyGluLeuAlaPheArgGlyGlyAspValValThrIleLeuGluAla 80
 Db CGCCCAAGAGGAGGAGCTGGCTCCGCAAGGGGAGCTGTACCATCTCTGGAGGCC 497
 Qy 81 CysGluAsnLysSerTrpTrpArgValLysHisThrSerGlyGlnGlyLeuLeu 100
 Db TGGCAGAACAGAGCTGTTACCGGCTCAGACACACACAGTGGAGAGGGGCTGCTG 557
 Qy 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 Db GCACCTGGGGGCGTGGCGGAGCGGAGGCGCTCCGAGACCCCAACCTCAGCCTCATG 617
 Qy 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
 Db CCGGGTCCACGGAGAAATCTCGGGCCAGAGGCTGCCAGACCTCAGCCTCCCGAG 677
 Qy 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysVal 160
 Db GATGGGCTGTTCTCGTGGTGGAGAGTCCCGCGGCCACCCCGGAGTACGTTCTGCTG 737
 Qy 161 SerPheGlyArgAspValIleHisLysTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
 Db AGCTTTGGCCGCGAGCTATCCACTACCGGCTGGTGGACCGCGAGGCTCAGCTCAGATC 797
 Qy 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisLysSerLysAspLys 200
 Db GATGAGGCGCTGTTCTCTGCACCTCATGAGCATGTGGAGCATTTACAGCAAGGACAG 857
 Qy 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
 Db GGGCGCTATCTGCACCAAGCTGTGAGACCAAGCGGAACACGGGAGCAAGTGGCGCAG 917
 Qy 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLysLeuThrLeuGlyAlaGln 240
 Db GAGGAGCTGGCCAGGGCGGCTGGTTACTGAACCTGCAGCATTTGACATTTGGAGCAAG 977
 Qy 241 IleGlyGluGlyGluPheGlyAlaValLeuGlnGlyLysTrpLeuGlyGlnLysValAla 260
 Db ATCGAGAGGAGAGATTGGAGCTGCTCGCAGGCTGAGTACCTGGGGCAAAAGGTGGCC 1037
 Qy 261 ValLysAsnIleLysCysAspValThrAlaGlnIlePheLeuAspGluThrAlaValMet 280
 Db GTGAAGATATCAAGTGTGATGTGACAGCCCAAGGCTTCTCTGGAGCAACAGCGGCTCATG 1097
 Qy 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
 Db ACGAAGATGCAACAGAGACCTGTGTCTCCGTGGCGGTGATCTCAGCAGGAGGCTG 1157
 Qy 301 TyrIleValMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
 Db TACATTTGCATGAGACAGTGTGAGCAAGGCAACCTGTGTGATCTTCTCGGAGCCGGGGT 1217

Qy 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMet 340
 Db CGAGCCCTCGTGAACACCGCTGCAGCTCTCTGATTTCTCTGCACGTCGGCGAGGGCATG 1277
 Qy 341 GluTrpLeuGlnSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
 Db GAGTACCTGGAGACCAAGAGCTGTGCACCGCACCTGGCGCCCGCCCAACATCTCTGTC 1337
 Qy 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuValLysAlaGluArgLysGly 380
 Db TCAGAGGAGCTGTGGCCAAAGGTGAGGACTTGTGGCTGGCCAAAGCGGAGGAGGG 1397
 Qy 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
 Db CTACACTCAAGCGCGCTGCCGTCAATGAGAGCGCGCCAGGCTGTCAAAACAGGGAAG 1457
 Qy 401 PheTrpSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrp 420
 Db TTCACACCAAGTGGATGTCTGGAGTTTGGGGTGTCTGTGGAGGTCCTTCTCATAT 1517
 Qy 421 GlyArgAlaProTrpTrpProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
 Db GGAGGGCTCCGTACCTTAATGTCACTGAAGAAGGTGTCTGGAGCCGTGGAGAGAGGG 1577
 Qy 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerCys 460
 Db TACCGCATGGAACCCCGGAGGGCTGTCCAGGCGCCGTCACGTCTCATAGCAGCTGC 1637
 Qy 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
 Db TGGGAGCAGAGCCCGCGCGGCGGCAACCTTCCGCAAACTGGCGGAGAGCTGGCCCG 1697
 Qy 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
 Db GAGCTACGCACTGAGGTGGCCCAAGCTCTGCTCAAGGGCAGAGCGGAGGCTCCAGC 1757
 Qy 501 SerProArgSerGlnGluPro 507
 Db TCGGCCGGAACCAAGAGGCC 1778

RESULT 3

US-09-977-261-1
 ; Sequence 1, Application US/09977261
 ; Publication No. US20030054527A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ULIRICH, AXEL
 ; APPLICANT: GISHIZKY, MIKHAIL
 ; APPLICANT: SURES, IRMINGARD
 ; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
 ; FILE REFERENCE: 038602/1259
 ; CURRENT APPLICATION NUMBER: US/09/977,261
 ; CURRENT FILING DATE: 2001-10-16
 ; PRIOR APPLICATION NUMBER: 08/232,545
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 2000
 ; TYPE: DNA
 ; ORGANISM: Unknown Organism
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (258)..(1778)
 ; FEATURE: (258)..(1778)
 ; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
 ; OTHER INFORMATION: kinase 1
 ; US-09-977-261-1

Alignment Scores:
 Pred. No.: 1.05e-266 Length: 2000
 Score: 2671.00 Matches: 507
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.008 Indels: 0
DB: 11 Gaps: 0
US-09-977-260-2 (1-507) x US-09-977-261-1 (1-2000)

QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
DB 258 ATGGCGGGGGAGAGCTCTGTGGTTCTCGCGGGCAATTCACGGCTGTGATTCTGTGGAG 317
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProAlaSerAla 40
DB 318 GAACCTCCCGGGTGAGCCCCCGCTTCCTCCAGCTGGACACCCCTCCCGCTAGCC 377
QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrIlyCysGluHisThr 60
DB 378 AGGATCCCAAGAGCGCTGGGCCCCCGGACACCACTGATACCAAAATGGCAGCACCC 437
QY 61 ArgProLysProGlyLysLeuAlaPheArgGlyGlyAspValValThrIleLeuGluAla 80
DB 438 CGCCCAAGCGAGGAGCTGGCTTCCTCCAGGGGAGCGTGGTCACTCCCTGGAGGCC 497
QY 81 CysGluAsnLysSerTrpArgValLysHisIleThrSerGlyGlnGlyLeuLeu 100
DB 498 TGGGAAACAGAGCTGTACCCCGCTCAAGCACACACAGTGGACAGAGGGGCTGCTG 557
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
DB 558 GCAGCTGGGGCGCTGGGAGCGGAGGCCCTCCGAGACCCCAAGCTCAGCTCAGC 617
QY 122 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
DB 618 CCGTGGTTCACGGAGAGTCTGGGCGAGAGGCTGTCCAGAGCTGACGCTCCCGAG 677
QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysVal 160
DB 678 GATGGGCTGTCTCGTGGGGAGTCCGGCGCCACCCCGGACATCGCTGCTGCTG 737
QY 161 SerPheGlyArgAspAlaIleHisIleTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
DB 738 AGCTTTGGCGCGACCTCATCCACTCCGCTGCTCACCAGCGCCACCTCAGCAATC 797
QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisIleSerIlySerLysAspLys 200
DB 798 GATGAGCGCGCTGTCTTCTTCGCAACCTCATGAGCATGTGGACATTCAGCAAGCAAG 857
QY 201 GlyAlaIleCysThrIlySerLeuValArgProLysArgLysHisGlyThrIlySerAlaGlu 220
DB 858 GGGCTATCTGCACCAAGCTGGTGAGACCAAGCGGAACAGCGGACCAAGTGGCGCAG 917
QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisIleuThrLeuGlyAlaGln 240
DB 918 GAGGACCTGGCGCGGGCGCTGTGTAACCTCAGCATTTGACATTGGGAGACAG 977
QY 241 IleGlyGluGlyValPheGlyAlaValLeuGlnGlyLysIleuThrLeuGlyLysValAla 260
DB 978 ATCGGAGAGGAGAGTGTGGAGCTGTCTGCAAGGGTGGAGACTGGGGCAAAAGGTGGCC 1037
QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGlyThrAlaValMet 280
DB 1038 GTGGAAGAATATCAAGTGTATGTAGACGGCCAGGCTTCTCTGACAGACAGCGCGTCAAG 1097
QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
DB 1098 ACGAAGATGCAACACAGAACTGGTGTCTCTGGGCTGATCTCTGACACAGGGGCTG 1157
QY 301 TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
DB 1158 TACATTGTATGAGGACGTGAGCAAGGGCAACCTGTGTAACCTTTCTGGGAGCCGGGCT 1217
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlnGlyMet 340
DB 1218 CGAGCGCTGTGTAACACCGCTCAGCTCTGAGATTCTCTGACAGTGGCCGAGGGCAGTC 1277

QY 341 GluTrpLeuGlnSerLysIlySerLeuValHisArgAspLeuAlaArgAsnIleLeuVal 360
DB 1278 GAGTACCTGGAGAGCAAGACTTGTGCAACCGGACCTGGCGCCCGCAACATCTGTGTC 1337
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
DB 1338 TCAGAGAGCTGTGGTGGCCAAAGTTCAGACGACTTGGCTGGCCAAAGCCGAGGAGGG 1397
QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
DB 1398 CTAGACTCAAGCGGCTGGCCGCTCAAGTGGAGCGGCGCCGAGGCTCTCAACACGGGAG 1457
QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrp 420
DB 1458 TTCACCAAGCAAGCGGATGTGTGGAGCTTTGGGGTCTCTCGGAGAGGCTTCTCATAT 1517
QY 421 GlyArgAlaProTrpProLysMetSerLeuLysGluValSerGluAlaValGlnLysGly 440
DB 1518 GAGCGGCTTCCTGACCTTAATGTCATGAAGAGTGTGGAGGCCCTGGAGAGGGG 1577
QY 441 TyrArgMetGlnProProGluLysCysProGlyProValHisValLeuMetSerSerCys 460
DB 1578 TACCGATGGAACCCCGCGAGGCTGTCCAGGCGCCGCTGCTCATAGACACTGC 1637
QY 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
DB 1638 TGGGAGCGAGAGCGCCCGCGGCCACCTTCGCAAACTGGCCGAGAAAGCTGGCGCG 1697
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
DB 1698 GAGCTAGGAGTGCAGAGTGGCCCGAGCTCGCTGTCAGGGCAGGACCGAGGCTCCACC 1757
QY 501 SerProArgSerGlnGluPro 507
DB 1758 TCGCCCGAAGCCAGAGGCC 1778

RESULT 4

US-10-084-817-341
; Sequence 341, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084, 817
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 341
; LENGTH: 1989
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Inocyte ID No. US20030119009A1 1794861CB1
US-10-084-817-341

Alignment Scores:

Pred. No.: 4,07e-261 Length: 1989
Score: 2617.00 Matches: 505
Percent Similarity: 99.618 Conservative: 0
Best Local Similarity: 99.618 Mismatches: 2
Query Match: 97.988 Indels: 2
DB: 14 Gaps: 0

US-09-977-260-2 (1-507) x US-10-084-817-341 (1-1989)

QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20


```

Db      |||ATGGGGGGCGAGGCTCTGTTCTCGGCGGCGATTTCACGGCTGATCTCTGAG 318
QY      21  GluLeuProArgValSerProArpPheLeuArpAlaTrpHisProProProValSerAla 40
Db      319  GAACTTCCCGGGGTAGCCCGCGCTCTCCGAGCGTGGCACCCCTCCGCTCCACCC 378
QY      41  ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGlnIsthr 60
Db      379  AGGATGCCACAGAGCGCTGGGCGCCCGGACCCAGTGTATCCAAATGCGAGCACACC 438
QY      61  ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValValThrIleLeuGlnAla 80
Db      439  CGCCCAAGCCAGGCGAGCTGGCTCCGCAAGGCGAGCTGTCCATCCATCCTGGAGGCC 498
QY      81  CysGluAsnLysSerTrpTrpArgValLysHisIsthrSerGlyGlnGluLeuLeu 100
Db      499  TCGGGAACAAAGAGTGTACCGCGCTCAAGCACACACCAAGTGGACAGAGGGCTGTG 558
QY      101  AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
Db      559  GCAGCTGGGGGCGCTCGGGAGGCGGAGGCCCTCTCCGAGACCCCAAGCTCAGCTCATG 618
QY      121  ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProGlu 140
Db      619  CCGTGTTCACAGGGAAGATCTCGGGCCAGAGCGCTGCCAGACGCTCAGCGCTCCGAG 678
QY      141  AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuGlyVal 160
Db      679  GATGGGCTGTTCCTGTGGTGGAGTCCGCGCGCCACCCCGGAGTACGCTGTGGCTG 738
QY      161  SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle 180
Db      739  AGCTTGGCCCGGAGAGTATCCACTACCGCGTGCACCGGAGCGGCACCTCACAATC 798
QY      181  AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTyrSerLysAspLys 200
Db      799  GATGAGCCCGTGTCTCTGCACCTCATGACATGTGAGCATTTACAGCAAGACAAAG 858
QY      201  GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
Db      859  GCGCTATCTGCACCAAGCTGGTGGACCAAGCGGAAACCGGAGCAAGTGGCGCAG 918
QY      221  GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisIleThrLeuGlyAlaGln 240
Db      919  GAGGAGCTGGCGGCGGGCGGTGTACTGAACCTGCAGCANTTACATTTGGAGACACAG 978
QY      241  IleGlyGluGlyGluPheGlyAlaValLeuGlnGlyGlyTrpLysGlyGlnLysValAla 260
Db      979  ATCGAGAGGAGGAGACTTGGAGCTCTCTCGCAGGGTGAATACCTGGGCAAAAGCTGCC 1038
QY      261  ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGlyThrAlaValMet 280
Db      1039  GTGAAGAAATATCACTGATGTGTGACAGCCCGGCTTCTCTGGAGAAAGCGGCTGTATG 1098
QY      281  ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
Db      1099  ACGAAGATGCAACAGACGACTGTGTGCTCCGCGGCGTGAATCTCTCAGCAGGCGCTG 1158
QY      301  TyrIleValIleMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
Db      1159  TACATTTCTCATGAGACAGTGTGAGCAAGGCAACCTGTGTAACCTTCTCGGAGCCGGGGT 1218
QY      321  ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLysHisValAlaGlnGlyMet 340
Db      1219  CGAGCCCTGCTGAACACGCTCAGCTCTGCGAGTTTCTCTGCACTGTGCGAGGCGATG 1278
QY      341  GluTrpLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
Db      1279  GAGTACCTGTGAGAGCAAGAGCTGTGACCGCGACCTGGCGCGCGCAACATCTGTGTC 1338
QY      361  SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380

```

```

Db      1339  TCAGAGACCTGCTGGCCCAAGGTCAGCGACTTTGGCTGCGCCAAAGCCGAGCGAAGGG 1398
QY      381  LeuAspSerSerArgLeuProValLysTrpTrpAlaProGluAlaLeuLysHisGlyLys 400
Db      1399  CTGACACTAAGCGGCTCCCGCTCAAGTGGAGCGGCGCGAGGCTCTCAAAACAGGGAAAG 1458
QY      401  PheTrpSerLysSerAspValTrpSerPheGlyValIleLeuLeuTrpGluValPheSerTyr 420
Db      1459  TTCACACAGCAAGTGGAGTGTCTGAGATTGTGGGGTGTCTGTGTGGAGAGTCTTCATAT 1518
QY      421  GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
Db      1519  GGACGGGCTCCGTACCTTAATAATGTCACTGAAGAAGAGGTGTGGAGGCGCTGGAGAAGGG 1578
QY      441  TyrArgMetGluProProGlnGlyGlyProGlyProValHisValLeuMetSerLys 460
Db      1579  TACCGATGGAACCCCGGAGGGCTGTCCAGGCGCCGTCGACAGTCTCATGTGAGCGTGC 1638
QY      461  TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
Db      1639  TGGAGAGCAGAGCGGCC-CGCGGCGCACCTTCCGCAAACTGGCCGAGAACGTCGCCGG 1697
QY      481  GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspLysSerThr 500
Db      1698  GAGCTACGACAGTGCAGGTGCCCAAGCTCCGCTCAGGGGACAGAGCGCGACGCTCC-ACC 1756
QY      501  SerProArgSerGlnGluPro 507
Db      1757  TCGCCCGAAGCGCAGAGGCC 1777

RESULT 5
US-09-954-531-188
; Sequence 188, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0.
; SEQ ID NO 188
; LENGTH: 2187
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-188

Alignment Scores:
Pred. No.: 3,27e-119 Length: 2187
Score: 1247.50 Matches: 238
Percent Similarity: 71.59% Conservative: 82
Best Local Similarity: 53.24% Mismatches: 118
Query Match: 46,718 Indels: 9
Gaps: 3

US-09-977-260-2 (1-507) x US-09-954-531-188 (1-2187)
QY      40  AlaArgMetProThrArg-----ArgTrpAlaProGlyThrGlnCys 53
Db      116  GCCAGAGCTCTGAGAGATGTACAGCAATACAGCGCGGCTGCGCATCCGTACAGAAATGT 175

```

QY 54 ILeThrIyScySgLnHstHrArGpProGlyGluLeuAlaPheAtrGlySGLyAsp 73
 Db 176 ATGGCCAAAGTACAACTTCCAGCGACTGCCAGAGAGACCTGCCCTTCTCCAAAGGAGAC 235
 QY 74 ValValThrIleLeuGluAlaIaGlyAsnIySerrTrpTyrArgValIySHstHstHr 93
 Db 236 GTGCTCACCATTGTGGCCGTCACCAAGACCCCACTGTTACAAAGCCAAAGCAAGAGTG 295
 QY 94 SerGlyGlnGluLeuLeuAlaIaGlyAlaLeuArgIuArgIuAlaLeuSerrAla 113
 Db 296 ---GGCCGTGAGGGCATCATCCAGCAACTACGTCCAGAGAGCGGAGCGGTGAAGCGC 352
 QY 114 AspProIySerrLeuMetProTrpPheHisGlyIySerrGlyGlnGluAlaVal 133
 Db 353 GGTAACCAACTCAGCTCATGCTGTGTTCCAGGGCAAGATCAACAGCGAGCGAGCTGAG 412
 QY 134 GlnGlnLeuGlnProGluAspGlyLeuPheLeuValIArgGluSerrAlaArgHstPro 153
 Db 413 CGGCTTCTGTACCCGGCGGAGACAGCCTTCTGTGGCGGAGAGCACCACACTACCC 472
 QY 154 GlyAspTyrValLeuGlyValSerPheGlyArgAspValIleHstTyrArgValLeuHis 173
 Db 473 GGAGACTACACGCTGTGCTGAGCTGCGAGCGCAAGGTGAGAGCATCCGACATCATGTAC 532
 QY 174 ArgAspGlyHstLeuThrIleAspGluAlaValPhePheGlyAsnLeuMetAspMetVal 193
 Db 533 CATGCCACGAGCTCAGCATCGAGAGAGAGGTGTACTTGAAGACCTCATGACAGCTGGTG 592
 QY 194 GlnHstTyrSerrIySAspIySGLyAlaIleCysThrIySLeuValArGpProIySArgIyS 213
 Db 593 GAGCAGTACACATCGACAGCGAGATGGACTGTACGGCGCTCATTTAAACCAAGGTGATG 652
 QY 214 HisGlyThrIySerrAlaGluGluLeuAlaArgAlaGlyTyrIleuLeuAsnLeuGln 233
 Db 653 GAGGGGACAGCTGGCGGGCGAGATGAGTTCTACCGCGAGCGGTGGCGCTGGAACATGAG 712
 QY 234 HisLeuThrIleuGlyAlaGlnIleGlyGluGlyGlnPheGlyAlaValIleuGlnGly 253
 Db 713 GAGCTGAGCTGTGTCGACAGCATTCGGAGAGGGAGTTCGAGACGGATGCTGGCGGAT 772
 QY 254 TyrLeuGlyGlnIySValAlaValIySAsnIleCysAspValIThrIaGlnAlaPhe 273
 Db 773 TACCGAGGGAACAAAGTCGCGCTCAAGTCAATTAAGACAGACCCCATGCGCCAGCCTTC 832
 QY 274 LeuAspGlyThrAlaValMetThrIySMetGlnHisGluAsnLeuValArGleuLeuGly 293
 Db 833 CTGGCTGAAGCTCAGTCATGACAGCAACTGCGCATGCAACTGGTGGTGCAGCTCTGGGC 892
 QY 294 ValIleuHstGln-----GlyLeuTyrIleValMetGlnHstValSerIySGLyAsn 311
 Db 893 GTATGCTGGAGAGAGAGGGCGGCTCTACATGCTCATGTAAGTACATGAGGCCAAGGGAGC 952
 QY 312 LeuValAsnPheLeuArgTrpArGlyArgAlaLeuValAsnThrIaGlnLeuGln 331
 Db 953 CTGTGTGACTACTGCGGTCTAGGGGTGGTCACTAGTGTGGCGGAGACTGTCTCTCAAG 1012
 QY 332 PheSerLeuHisValaIaGluGlyMetGluTyrLeuGluSerIySLeuValIHisArg 351
 Db 1013 TTCTCCCTAGATGTCTGCGAGGCGCATGATACCTGGAGGCAACAATTCTCGTGCATCGA 1072
 QY 352 AspLeuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaIySValSerAspPhe 371
 Db 1073 GACTGTGCTGCGCCGCAATGTGTGTGTGAGAGCAAACTGGCCAAAGGTGCAGCTTTT 1132
 QY 372 GlyLeuAlaAlaArgIuArgIySGLyLeuAspSerrArgLeuProValIySTPTr 391
 Db 1133 GGTCTACCAAGAGAGCGCTCCAGCAGCAGACAGCGCAAGCTGCCAGTCAAGTGGACA 1192
 QY 392 AlaProGluAlaLeuIySGLySpherHstIySerrIySAspValITrpSerrPheGly 411
 Db 1193 GCCCTGAGGCGCTGAGAGAGAGAAATCTCCACTAAGTCTGACGCTGGAATTTGCGA 1252
 QY 412 ValLeuLeuTrpGluValPheSerTyrGlyArgAlaProTyrProIySMetSerLeuS 431

Db 1253 ATCCTTCTGGGAAATCACTACTTTGGCGAGTGCCTTATCCAAATTCCTCCGAG 1312
 QY 432 GluValSerGluAlaValIySGLyTyrArgMetGluProGluGlyCysProGly 451
 Db 1313 GACTGTCTCCCTCGGGTGGAGAGAGGTGTACAAATGTATGCCCCAGCGGCTGCCGCC 1372
 QY 452 ProValHstValLeuMetSerrCysTrpGluAlaIuProAlaArgArGpProPhe 471
 Db 1373 GCAGTCTATAGATCAAGAACTGCTGGCACCTGGAGCGCCCATGCGGCTCTCTTC 1432
 QY 472 ArgIySLeuAlaGluIySLeu 478
 Db 1433 CTACAGCTCCGAGAGCAGCTT 1453
 RESULT 6
 US-10-298-377A-1
 : Sequence 1, Application US/10298377A
 : Publication No. US20030130209A1
 : GENERAL INFORMATION:
 : APPLICANT: The Scripps Research Institute
 : APPLICANT: Cheresch, David A.
 : APPLICANT: Paul, Robert
 : APPLICANT: Elliceiri, Brian
 : TITLE OF INVENTION: Method of Treatment of Myocardial
 : FILE REFERENCE: Infarction
 : CURRENT APPLICATION NUMBER: US/10/298, 377A
 : CURRENT FILING DATE: 2002-11-18
 : PRIOR APPLICATION NUMBER: 10/298, 377
 : PRIOR FILING DATE: 2002-11-18
 : PRIOR APPLICATION NUMBER: 09/470, 881
 : PRIOR FILING DATE: 1999-12-22
 : PRIOR APPLICATION NUMBER: 09/538, 248
 : PRIOR FILING DATE: 2000-03-29
 : PRIOR APPLICATION NUMBER: PCT/US99/11780
 : PRIOR FILING DATE: 1999-05-28
 : PRIOR APPLICATION NUMBER: 60/087, 220
 : PRIOR FILING DATE: 1998-05-29
 : NUMBER OF SEQ ID NOS: 4
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 1
 : LENGTH: 2187
 : TYPE: DNA
 : ORGANISM: homo sapiens
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (134)...(1486)
 : US-10-298-377A-1
 Alignment Scores:
 Pred. No.: 3,27e-119 Length: 2187
 Score: 1247.50 Matches: 238
 Percent Similarity: 71.59% Conservative: 82
 Best Local Similarity: 53.24% Mismatches: 118
 Query Match: 46.71% Indels: 9
 Gaps: 3
 US-09-977-260-2 (1-507) x US-10-298-377A-1 (1-2187)
 QY 40 AlaArgMetProThrArg-----ArgTrpAlaProGlyThrGlnCys 53
 Db 116 GCCAGAGTCTCTAGAGAAGATGTACGAATACAGCGCCGCTTGTGCAAGAGATGT 175
 QY 54 ILeThrIyScySgLnHstHrArGpProGlyGluLeuAlaPheAtrGlySGLyAsp 73
 Db 176 ATGGCCAAAGTACAACTTCCAGCGACTGCCAGAGAGACTGCTTGTGCAAGAGAGAC 235
 QY 74 ValValThrIleLeuGluAlaIaGlyAsnIySerrTrpTyrArgValIySHstHstHr 93
 Db 236 GTGCTCACCATTGTGGCCGTCACCAAGACCCCACTGTTACAAAGCCAAAGCAAGGTG 295
 QY 94 SerGlyGlnGluLeuLeuAlaIaGlyAlaLeuArgIuArgIuAlaLeuSerrAla 113

```
Db      296 ---GGCGCTGAGGCGCATCCAGCACTGACCTGCAAGAGCGGAGCGCTGAGGCG 352
Qy      114 ASPRQYLSLEuSeLleuMeTTrPpheNisGLyLSIleSeGLyGLuAlaVal 133
Db      353 GGTACCAACACTGAGCTCATGCTTGGTTCACGGCAAGATCACAGGAGCGAGCTGAG 412
Qy      134 GLuNleuGLuNProGLuNAPRleuNpheuValArgGLuSeLAlaArgHisPro 153
Db      413 CGGCTTCTGTACCCCGGAGACAGGCTGTTCTGTGTCGGGAAGACCACTACACCC 472
Qy      154 GLyAPRyValLeuCySValSerPheGLyARgAPRValIleHisTyARgValLeuHis 173
Db      473 GGAGACTACAGCTGTGGGTGAGTGCAGCGCAAGGTGAGACTACAGCATCATGTAC 532
Qy      174 ARGAPRGLyNleuThrlleAsPRGLuAlaAlPhePheCySAsNleuMeLAsPmeVal 193
Db      533 CATGCCACCAACTGAGATGCAGAGGAGGTGACTTGGACAACTTCATGCAGCTGTG 592
Qy      194 GLuHISTySerLySAPRySGLyAlaIleCySThrlYLeuValArgProLySArgLyS 213
Db      593 GAGCACTACACTCAGACGCAAGATGACTCTGTACGCGCTCATTAACCAAGGTCATG 652
Qy      214 HISGLyThrlYSerAlaGLuGLuGLuLeuAlaArgAlaGLyTrPLeuLeuAsNleuGLN 233
Db      653 GAGGCGACAGTGGCGCCAGAGTATGTACCGCAGCGCGCTGGCGCCCTGAACATGAG 712
Qy      234 HISleuThrlleuGLyAlaGLuNleuGLyGLuGLyPheGLyAlaValleuNleuGLy 253
Db      713 GAGCTGAAGCTCTCTCAGACACATCGGGAAGGGGAGTTCGAGCTGTGCTGGCGCT 772
Qy      254 TyTleuGLyGLuNleuValAlaValAlaLySAsNleuLySAsAPRValThrlleuNleuPhe 273
Db      773 TACCGAGGAGACAAAGTGGCGCTCAAGTGCATTAAGAACAGACGCCACTGCCAGGCTTC 832
Qy      274 LeuAPRGLyThrlAlaValleuThrlYSerMetGLNHisGLuAsNleuValArgLeuGLy 293
Db      833 CTGGCTGAAGCTCTCAGTATGACGACACTGCGCATAGACACTGCTGAGCTCTGGGCG 892
Qy      294 ValIleleuNisGLN-----GLyLeuTyTleuValMetGLNHisValSerLySGLyAsN 311
Db      893 GTAGCTGTGTGAGAAAGGCGGCTCATCATCTGACATCGACTGATCATGATGCGCAAGGGAGAC 952
Qy      312 LeuValAsNpheLeuArgThrlArgGLyAlaArgAlaValAsNThrlleuNleuGLN 331
Db      953 CTGTGGACTACTCTCGGCTGAGGCTGCGTCACTGCTGGCGGAGACACTGCTCTCTCAG 1012
Qy      332 PheSerLeuNisValAlaGLuGLyMetGLyTyTleuGLySerLySLeuValHisArg 351
Db      1013 TTTCTGCTAGATGTCTGCGAGGCGCATGTGAACTGAGAGGCGCAACAATTTCTGCATGCA 1072
Qy      352 AsPLeuAlaAlaArgAsNleuValSerGLuAsPLeuValAlaLySAlaSerPhe 371
Db      1073 GACCTGTGGTCCCGGAAATGTCTGTGTGAGCAACACTGAGGCGCAAGGTCAGCACTTT 1132
Qy      372 GLyLeuAlaLySAlaGLuArgLySGLyLeuAsPLeuSerArgLeuProValLySTrPTr 391
Db      1133 GGTCTCACCAAGAGGCGGTCCAGCAACCGACAGGCGGCAACTGCGACTCAAGTGGACA 1192
Qy      392 AlaPRGLuAlaLeuLySHisGLySpherThrlYSerLySAsPValTrPserPheGLy 411
Db      1193 GCCCTGAGGCGCTGAGAGAGAAATTTCTCACTAAGTGTGAGCTGTGAGATTTCGGA 1252
Qy      412 ValIleuLeuTrPGLuValPheSerTyTleuArgAlaProTyTleuMetSerLeuLyS 431
Db      1253 ATCTCTCTGTGGAAATCTACTCTTGGCGAGTGCCTTATCCAAAGTTCTCCCTGAG 1312
Qy      432 GLyValSerGLuAlaValGLySGLyTyTleuMetGLyProGLuGLyLysProGLy 451
Db      1313 GACGTCTGCTCCGCTGAGAGAGGCTCAAGATGAGAGCCCGGAGCGCTGCGCGCC 1372
Qy      452 ProValHisValleuMetSerCySTrPGLuAlaGLuProAlaArgTrPProPhe 471
```

```
Db      1373 GCAGTCTGATGATCATGAGAACTGTGGCACTGAGCGCGGCATGGCGCTCTTC 1432
Qy      472 ArgLySleuAlaGLyLysleu 478
Db      1433 CTACAGCTCCGAGAGCAGCTT 1453

RESULT 7
US-10-177-293-87
; Sequence 87, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatz, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, VIC
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Puzsai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 2420
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-87

Alignment Scores:
Pred. No.: 3,76e-119 Length: 2420
Score: 1247.50 Matches: 238
Percent Similarity: 71.598 Conservative: 82
Best Local Similarity: 53.248 Mismatches: 118
Query Match: 46.718 Indels: 9
DB: 14 Gaps: 3

US-09-977-260-2 (1-507) x US-10-177-293-87 (1-2420)

Qy      40 AlaArgMetProThArg-----ArgTrPAlaProGLyThrlNleuCyS 53
Db      395 GCCAGAGCTCTGAGAAAGATGTGCAGCAATACAGCGCGCTGCGCAATCCGTCAGAGATGT 454
Qy      54 IleThrlYScyGLuNHisThrlArgProLySGLyLleuAlaPheArgLySGLyAsP 73
Db      455 ATTGCCAACTACACTTCCAGCGCACTGCGGAGCAAGACTGCGCTCTGCAAGAGAGAC 514
Qy      74 ValValThrlleuGLuAlaCySGLuNlySsetTrTyTleuArgVallyNHisNThr 93
```

```
Db 515 GTGCTCAACATTTGGCCGTCACACAGAGCCCAAGTGGTCAACAAACAAACAGAGTG 574
Qy 94 SerIyGlnGluGlyLeuLeuAlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAla 113
Db 575 ---GGCCGTGGAGGATCATCCAGCAACATCACTCCAGAGCGGAGGCGGTGAAGCGC 631
Qy 114 AspProLysSerLeuMetProTyrPheHisGlyLysLysSerGlyGlnGluAlaVal 133
Db 632 GGTACCAAACTCAGCTCATGCTTGGTTCACGGCAAGATCAACACGGGAGCAGCTGAG 691
Qy 134 GlnGlnLeuGlnProProGluAspGlyLeuPheLeuValArgGluSerAlaArgHisPro 153
Db 692 CGAGCTCTGTACCCGGGAGACAGGCTGTCTGTGGCGGAGACACCACTACCC 751
Qy 154 GlyAspTyrValLeuCysValSerPheGlyArgAspValIleHisTyrArgValLeuHis 173
Db 752 GGAGACTACACAGCTGTGCTGAGCTGCAGCGCAAGGTGAGACCTACCATCATGATATC 811
Qy 174 ArgAspGlyHisLeuThrIleAspGlyAlaValPhePheCysAsnLeuMetAspMetVal 193
Db 812 CATGCCACGAACTCAGCATCGACGAGAGAGGTGTACTTTGAGAACTCATATGCAAGTGTG 871
Qy 194 GluHisTyrSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLysArgLys 213
Db 872 GAGACTACACACTCAGACGAGATGAGATGCTGTACGGCCCTCATTAACCAAGTGTATG 931
Qy 214 HisGlyThrLysSerAlaGluGluGluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGln 233
Db 932 GAGGGCACAGTGGCGGCGGAGATGATTTACCGCGAGCGGCGGCGCTGACATGATGAG 991
Qy 234 HisLeuThrLeuGlyAlaGlnIleGlyGluGlyGlnPheGlyAlaValLeuGlnGlyGlu 253
Db 992 GAGCTGAAGTGTCTGCAGACCATCGGAGAGGGAGTTCGGAACAGTATGCTGGCGGAT 1051
Qy 254 TyrLeuGlyGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGlnAlaPhe 273
Db 1052 TACCGAGGAAACAAAGTCGCGCTCAAGTCAATTAAAGACACCCCACTGCCAGGCTTC 1111
Qy 274 LeuAspGluThrAlaValMetThrLysMetGlnHisGlnAsnLeuValArgLeuGly 293
Db 1112 CTGGCTGAAGCTCAGTACGAGCAACTCGGCATGACAACTGGTGAGCTCTGGGC 1171
Qy 294 ValIleLeuHisGln-----GlyLeuTyrIleValMetGlnHisLysSerGlyAsn 311
Db 1172 GTATGCTGAGGAGGAAGGCGGCTCTACATCTACTGATGATGCGCAAGGGGAGGC 1231
Qy 312 LeuValAsnPheLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuGln 331
Db 1232 CTGTGAGACTACTGGGCTGTAGGGTCTGCTGCTGCGGAGACACTCTCTCTCAAG 1291
Qy 332 PheSerLeuHisValAlaGluGlyMetGluTyrLeuGluSerLysLysLeuValHisArg 351
Db 1292 TTCCTCCATGATTCGCGAGGCGCATGATACCTGAGGCAACAATTCGTGATCGA 1351
Qy 352 AspLeuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaLysValSerAspPhe 371
Db 1352 GACCTGGCTGCCCGCATGCTGCTGTGAGGACAAGCTGGCCAAAGGTCACGCACTTT 1411
Qy 372 GlyLeuAlaLysAlaGluArgLysGlyLeuAspSerSerArgLeuProValLysTyrPhe 391
Db 1412 GGCTCTCACCAGAGGCGTCCAGCACCGACAGGAGGAGAGCTGCATCAAGTGGACA 1471
Qy 392 AlaProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspValTyrSerPheGly 411
Db 1472 GCCCTGAGGCGCTGAGAGAGAAGAAATTCACATAACTGACTGTGGAGTTTCGGA 1531
Qy 412 ValLeuLeuTyrGluValPheSerTyrGlyArgAlaProTyrProLysMetSerLeuLys 431
Db 1532 ATCCTTCTCTGGAAATCTACTCTTGGCGAGTGCCTTATTCACAAAGATTCCTCCGAG 1591
Qy 432 GluValSerGluAlaValGluLysGlyTyrArgMetGluProProGluGlyCysProGly 451
```

```
Db 1592 GACGTGCTCCCTCGCGGTGGAGAGGCTACAAAGATGATGCCCCGACGCTCCCGCC 1651
Qy 452 ProValHisValLeuMetSerCysTyrPheGluAlaGluProAlaArgArgProPhe 471
Db 1652 GCACGTATAGATCATGATGAAGATGCTGTGGCACTGAGACGGCGGCATGCGGCTCTC 1711
Qy 472 ArgLysLeuAlaGluLysLeu 478
Db 1712 CTACAGCTCCGAGAGCAGCTT 1732

RESULT 8
US-09-954-456-1983
Sequence 1983, Application US/09954456
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
TITLE OF INVENTION: Sets
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: Patent version 3.0
SEQ ID NO 1983
LENGTH: 2015
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-1983

Alignment Scores:
Pred. No.: 3,4e-65 Length: 2015
Score: 725.50 Matches: 180
Percent Similarity: 51.25% Conservative: 87
Best Local Similarity: 34.55% Mismatches: 187
Query Match: 27,168 Gaps: 67
DB: 10 Gaps: 15

US-09-977-260-2 (1-507) x US-09-954-456-1983 (1-2015)
Qy 34 HisProProValSerAlaArg----- 41
Db 272 CATCCACCA-----TCAAGCCGGGCGCTTAATAGCCACACAGCAACACAGCAATCA 325
Qy 42 -----MetProThrArgTArgTArgTArgTArgTArgTArgTArgTArgTArgTArgT 57
Db 326 GGGAGGACGAGCTGTAGAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 364
Qy 58 GluHisThrArgPro-LysPro-GlyGluLeuAlaPheArgLysGlyAspValValThrI 77
Db 365 ATGATTACAGGAGCCATTCACACGAAAGACCTCAGCTTCAGAAAGGGGAGACAGATGCTG 424
Qy 77 IeLeuGluAlaCysGluAsnLysSerTyrTyrArgValLysHisThrSerGlyGln 97
```


Db 581 TCCTCCGAGACATGCTGGCTCTTCATGATCCGGAGTACGAGACCACTAAAGAA 640
Qy 155 spyrValLeuCyVal-----SerpheGlyArgAspValIleHisTyra 170
Db 641 GCTACTCTTTTGTCCGTGGAGACTACAGACCTCCGGCAGGAGATACCGTGAAACATTACA 700
Qy 170 rgyValLeuHisArgAsp---GlyHisLeuThrIleAspGluAlaValPhePheCysAsnL 189
Db 701 AGATCCGAGACCTTCGACACGCGGGCTTCTACATATACCCCGAAGCACTTCAGCACTC 760
Qy 189 eumEtaSpmetValGluHisTyreSerLysAspLysGlyAlaIleCysThrLysLeuVal 209
Db 761 TCGAGAGCTGGTGGACACTACAGAGGAGGAGACGAGGCTCTGCGAAGAACTGTGCG 820
Qy 209 rgrProLysArgLysHisGlyLysSerLeuGluGluLysLeuAlaArgLysLysLys 229
Db 821 TGCCCTGCATG-----TCTTCCAGGCCCGAAGGAGCTTGGGAGAAAGATCCCTGGG 871
Qy 229 euleuAsnLeuGluHisLeuThrLeuGlyAlaGluIleGlyGluGlyLysPheGlyAla 249
Db 872 AGATCCCTCGGAGATCCCTCAAGCTGGAGAGAAACTTGGAGCTGGGAGCTTGGGAGAG 931
Qy 249 alleuGluGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 267
Db 932 TCTGGATGGCCACTACACAGACGACACAGGCTGGAGTGAAGCATGAGGAGGAGGAG 991
Qy 267 spValThrAlaGluAlaPheLeuAspGluThrAlaValMetThrLysMetGluHisLys 287
Db 992 GCATGTGGTGGAGGCTCTCTCGCAGAGGCAAGCGATGAAAGCTTCAGCATGATACA 1051
Qy 287 snLeuValArgLeuLeuGlyAlaIleLeuHisGluGlyLysLysLysLysLysLysLys 307
Db 1052 AGCTGTCAAACTTCATGCGGTGGTCCAGAGGCCATCTGATCATCAGGAGTTC 1111
Qy 307 alSerLysGlyAsnLeuValAsnPheLeuArgThrArgGlyArgAlaLeuValAsnThr 327
Db 1112 TGGCCAAAGAGAGCTTCTGCTGACTTCTGAAAAGTGTAGAGGAGGAGGAGGAGGAG 1171
Qy 327 lacLeuLeuGluGlnPheSerLeuHisValAlaGluGlyMetGluTyrlLeuGluSerLys 347
Db 1172 CAAGACTCATGACTTCTGACCCAGATGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1231
Qy 347 yslLeuValHisArgAspLeuAlaAlaArgAsnIleLeuValSerGluAspLeuValAla 367
Db 1232 ACTACATCCACGAGACTCCGAGCTGCCAACAATCTGTCTGCTGCTGCTGCTGCTG 1291
Qy 367 yslSerAspPheGlyLeuAlaLysAla-----GluArgLysGlyLeuAsp 383
Db 1292 AGATTCCTGCTGCTTGGCTGGCCGCGGTGCTAGAGGAGGAGGAGGAGGAGGAGGAG 1351
Qy 383 erSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLysPheThr 403
Db 1352 GGGCCAAAGTTCCTCATCAATGAGAGCTCTGTAAGGAGGAGGAGGAGGAGGAGGAGGAG 1411
Qy 403 erLysSerAspValLysPheGlyValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 423
Db 1412 TCAGTCTAGAGCTGTGCTCTTGTGATCTGCTGATGAGATCGTCACTAGGCGCGGA 1471
Qy 423 labProLysProLysMetSerLeuLysGluValSerGluAlaValGluLysGlyTyrlArg 443
Db 1472 TCCCTTACCCAGGAGATGTCAAAACCTGAATGATCCGAGCTTCGAGACGCTGATCCGGA 1531
Qy 443 eteLysProLysGlyCysProGlyProValHisValLeuMetSerSerCysTrpGlu 463
Db 1532 TGCCCTGCCAGAGACTGCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1591
Qy 463 lagLysProLysArgProLysPheArgLysLeuAlaGluLysLys----- 478
Db 1592 ACCGTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1651
Qy 479 --AlaArgGluLeuArgSerAla-----GlyAlaProLysSerValSerG 493

Db 1652 CGGCCACAGAGAGCCACTACCAACAGACGACCATGATGAGGAGGAGGAGGAGGAGG 1711
Qy 493 lylGlnAspAlaAspGlySer-----ThrSerProArgSerGlnGluPro 507
Db 1712 GGTGCCACAGGTGGTGGCTTCAGAGTGGCTCCAGACCATTCGCCAGGAGGAGGAGGAGGAG 1770
RESULT 10
US-09-917-800A-1611
Sequence 1611, Application US/09917800A
Patent No. US20020119462A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Castle, Arthur
APPLICANT: Elashoff, Michael
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Molecular Toxicology Modelling
FILE REFERENCE: 44921-5038-US
CURRENT APPLICATION NUMBER: US/09/917, 800A
CURRENT FILING DATE: 2001-07-31
PRIORITY APPLICATION NUMBER: US 60/222,040
PRIORITY FILING DATE: 2000-07-31
PRIORITY APPLICATION NUMBER: US 60/222,880
PRIORITY FILING DATE: 2000-11-02
PRIORITY APPLICATION NUMBER: US 60/290,029
PRIORITY FILING DATE: 2001-05-11
PRIORITY APPLICATION NUMBER: US 60/290,645
PRIORITY FILING DATE: 2001-05-15
PRIORITY APPLICATION NUMBER: US 60/292,336
PRIORITY FILING DATE: 2001-05-22
PRIORITY APPLICATION NUMBER: US 60/295,798
PRIORITY FILING DATE: 2001-06-06
PRIORITY APPLICATION NUMBER: US 60/297,457
PRIORITY FILING DATE: 2001-06-13
PRIORITY APPLICATION NUMBER: US 60/298,884
PRIORITY FILING DATE: 2001-06-19
PRIORITY APPLICATION NUMBER: US 60/303,459
PRIORITY FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 1740
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1611
LENGTH: 1911
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_013185
US-09-917-800A-1611
Alignment Scores:
Pred. No.: 3,43e-64 Length: 1911
Score: 715.50 Matches: 172
Percent Similarity: 54.01% Conservative: 84
Best Local Similarity: 36.29% Mismatches: 179
Query Match: 26.79% Indels: 39
DB: 10 Gaps: 13
US-09-977-260-2 (1-507) x US-09-917-800A-1611 (1-1911)
Qy 27 ProArgPheLeuArgLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 46
Db 281 CACGTCCTCCCTTAAAGAGCTGG---GACCGAAGCATTCACAGCCTGCCCCGGGTTGG 337
Qy 47 Trp---AlaProGlyThrGlnCysIleThrLysCysGluHisThrArgPro-LysPro-G 65
Db 338 TGGAGGCTCTGTGAGACACCATGCTGTGCTGCACTGTACAGCATGTAGGCCATTACCGTG 397
Qy 65 lylGluLeuAlaPheArgLysGlyAspValValThrIleLeuGluAlaCysGluAsnLys 85
Db 398 AAGACCTCAGCTTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 451
Qy 85 erTrpTyrlArgValLysHisHisThrSerGlyGlnGluGlyLeuLeuAlaValAla 105

Db 452 AGTGTGGAGAGCCGTTCCCTGGCTACCAAGAAAGAGGCTATATCCCAACATATATG 511
Qy 105 euaTggluaTggluaLeuSerAlaSpProLysLeuSerLeuMetProTriphHisG 125
Db 512 TAGCTCGAGTTAACTCTTTGGAGACTGAG-----GAGTGGTTCTTCA 553
Qy 125 lYlYsIleSerGlyngInuAlaValngInuLeuInProProGluuAsp-----GlyL 143
Db 554 AGGATATCAGCCGAGAGATGCGAGGCCACCTGCGTCCCGGAGAACATGCTGGCT 613
Qy 143 eupHeleuValaTggluSerAlaArgHisProGlyAspTyrValLeuGlyVal----- 160
Db 614 CCTTATGATCTCCGGACAGCTAGACCCAAAGGAGAGCTACTCTTCTGTGAGACT 673
Qy 161 -----SerPheGlyAlaGspValIleHisTyrArgValLeuHisArgsp---GlyH 177
Db 674 TTGACCCCGACAGGAGACAGCGGTGAAGCATTTAAATCCGACACTGGACAGTGGAG 733
Qy 177 lSleuThrIleAspGluAlaValaPhePheCysAsnLeuMetAspMetValGluHisTyrS 197
Db 734 GGTCTACATCTCTCCGAGAGACACTTCACAGCCTGCAGGAAGTCTGCTCCACTCA 793
Qy 197 erLYsAspLYsGLYAlaIleCysThrLYsLeu-----ValaArgProLYs---A 212
Db 794 AGAAGGGAGAGATGGCTGTGCGAAGAGCTGACAGTCCCTGTGTCTCCGAAGCC 853
Qy 212 rGLYSHISGLYThLYsSerAlaGluGluGluAlaArgAlaGlyIlePheLeuAsnL 232
Db 854 AGAAGCCATGGAGAAAGATGCC-----TGGAGATTCCTC 889
Qy 232 eugInHisLeuThrLeuGlyAlaGluInleGlyGluGlyPheGlyAlaValaLeuGluG 252
Db 890 GAGATCCCTCTCAGATGAGAAAGAACTGGAGCCGCACTTTGGAGAACTGTGGATG 949
Qy 252 lYgluTyr---LeuGlyInLYsValaValaLYsAsnIleLYs---CysAspValIle 270
Db 950 CCACCTACACAAACACACAAAGTGGCGGTGAAGACAAATGAAGCAGAGGAGCATGTCTG 1009
Qy 270 lAglInAlaPheLeuAspGluThrAlaValMetThrLYsMetGluHisGluAsnLeuVala 290
Db 1010 TGGAGGCTCTCTGGCAGAGGCCCACTGATGAAGACCTTACAGCATGATTAACCTGGGA 1069
Qy 290 rGLeuleuGlyValIleLeuHisGluInLYsTyrIleValaMetGluHisValSerLYsG 310
Db 1070 AGCTAACAGCTGTGTCTCTCAGAGGCCCATCTTATGTACCGAGTTCATGGCCAAAG 1129
Qy 310 lYAsnLeuValaAspPheLeuArgThrArgAlaLeuValaAsnThrAlaGluLeuL 330
Db 1130 GAAGCTCTGTGACTTCTCTCAAGAGTGAAGAGCAGCAAGCAGCAGCTGCCAAACTCA 1189
Qy 330 eugInPheSerLeuHisValaIleGluGlyMetGluTyrLeuGluSerLYsLYsLeuValH 350
Db 1190 TTGACTTCTCAGCCAGATTTCAAGAGGCAATGCTTTCATTGACAGAGGAATACATACC 1249
Qy 350 lAsArgAspLeuAlaAlaArgAsnIleLeuValSerGluAspLeuValaAlaValaSerA 370
Db 1250 ACCGAGACCTCCGGGCTGCCAACAATCTTGTCTGCATCAGTGTGTGAATCGCTG 1309
Qy 370 sPrPheGlyLeuAlaLYsAla-----GluAqLYsGLYLeuAspSerSerArgL 386
Db 1310 ACTTTGGACTGGCAGGATCATCGAGCAATGAGTACACAGCTCGGAGAGAGCCCAAGT 1369
Qy 386 eupProVALYsTrPheAlaProGluAlaLeuLYsSHISGLYsPheThrSerLYsSerA 406
Db 1370 TCCCATATCAAGTGCAGACCTCTGAAGCCATCACTTGGCTCTTCACTCAAGTAG 1429
Qy 406 sPVALYsTrPheGlyValaLeuLeuTyrGluValaPheSerTyrGlyArgAlaProTyrP 426
Db 1430 ATGCTGGTCTTGGTATCTCTGATGAAATCGTCACTACAGCGCGGATCCCTTACC 1489
Qy 426 rOLYsMetSerLeuLYsGLYValaSerGluAlaValaGluLYsGLYTYrArgMetGluProP 446
Db 1490 CAGTATGTCAAAACCCAGAGGTGATTCGAGCACTAGACATGAGGTACCGTATGCTCGAC 1549

Qy 446 rOLuGLYCySProGLYProVALHISValleuMetSerCysTriPGLuAlaGluProA 466
Db 1550 CAGATTAAGTGCAGAGAGAGCTTACACTATATGATATCGCTGCTGGAAGAACCTGCAG 1609
Qy 466 lAaArgArgProPheArgLYsLeuAlaGluLYsLeu 478
Db 1610 AGGAAGGCGCCACTTTCGATATACATCCAGACGTCGTG 1647
RESULT 11
US-10-298-377A-3
Sequence 3, Application US/10298377A
Publication No. US20030130209A1
GENERAL INFORMATION:
APPLICANT: The Scripps Research Institute
APPLICANT: Cheresih, David A.
APPLICANT: Paul, Robert
APPLICANT: Elicelrli, Brian
TITLE OF INVENTION: Method of Treatment of Myocardial
FILE REFERENCE: TSRI-651-5
CURRENT APPLICATION NUMBER: US/10/298,377A
CURRENT FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: 10/298,377
PRIOR FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: 09/470,881
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: 09/538,248
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: PCT/US99/11780
PRIOR FILING DATE: 1999-05-28
PRIOR APPLICATION NUMBER: 60/087,220
PRIOR FILING DATE: 1998-05-29
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 4517
TYPE: DNA
ORGANISM: homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (208)...(1839)
US-10-298-377A-3
Alignment Scores:
Pred. No.: 4,21e-63 Length: 4517
Score: 710.00 Matches: 163
Percent Similarity: 54.61% Conservative: 86
Best Local Similarity: 35.75% Mismatches: 165
Query Match: 26.58% Indels: 42
DB: 15 Gaps: 11
US-09-977-260-2 (1-507) x US-10-298-377A-3 (1-4517)
Qy 66 GluLeuAlaPheArgLYsGLYAspValaThrIleLeuGluAlaCysGluAsnLYsSer 85
Db 532 GACTTCATTTAAAGAGGTAAGATTTCAATAATTAACAATACATACGGAAGGA---GAT 588
Qy 86 TrPTyrArgValLYsSHISHisThrSerGlyngInuLeuAlaIleGlyAlaLeu 105
Db 589 TGTGTGGAGCAAGATCAATCGCTACAGAAAGATGTATATCCGCAAAATATATGTA 648
Qy 106 ArgGluArgGluAlaLeuSerAlaAspProLYsLeuSerLeuMetProTriphHisGLY 125
Db 649 GCGGCTGCAGATTCATTCACGCAAGAA-----GAATGATATTTTGGC 690
Qy 126 lYsIleSerGlyngInuAlaValaGluInuLeuInProPro-----GluAspGlyLeu 143
Db 691 AAATGGGGAAGAAAGATGCTGAAGATTAATCTTTGAATCTCGGAATCAACAGAGTATT 750
Qy 144 PheLeuValaArgGluSerAlaArgHisProGlyAspTyrValaLeuCysValSer----- 161
Db 751 TTCTTAGTAAGAGAGAGTGAAGAAACAATGAAGTGTCTTATTCCTTTGATTCGTGATTTGG 810

OY 162 -----PheGlyArgAspValIleHisThrArgValLeuHisAsp---GlyHis 177
 DB 811 GATGAGATTAAGGGGTGCAATGTAAACCTACAAATTTGGAACCTTGCAATGCTGGA 870
 OY 178 LeuThrIleAspGluAlaValIlePheMetCysAsnLeuMetValGluHisTyrSer 197
 DB 871 TACTATATCAACACCAAGCAACATTTGATCTGTGCAAGAAATTTGGTGAACCTACACA 930
 OY 198 LysAspLysGluAlaIleCysThrLysLeu-----ValArgProLysArgLysHisGly 215
 DB 931 GAACATGCTGATGTGTTATGCTCCAAAGTGTGCAACCTGTGTGTCACACTGTGAACCTGAG 990
 OY 216 ThrLysSerAlaGluGluGluAlaArgAlaGlyTyrProLeuAsnLeuGluHisLeu 235
 DB 991 ACTCAAGT-----CTAGCAAAAGATGCTTGGGAATCCCTGAGAACTTTTG 1038
 OY 236 ThrLeuGluAlaGluIleGluGluGluPheGluValAlaValLeuGluIleGluTyrLeu 255
 DB 1039 CGACTAGAGGTTAACTAGACAAAGATGTTTGGCGCAACTGTGCGATGGGAACTGGAAAT 1098
 OY 256 Gly----GlnLysValAlaValLysAsnIleLys---CysAspValThrAlaGluAlaPhe 273
 DB 1099 GGAACSCAGAAAGTAGCAATCAAAACACATCAACAGTACATGATGCCAGAACTTTTC 1158
 OY 274 LeuAspGluThrAlaValMetThrLysMetGlnHisGluAsnLeuValAlaGluLeuGly 293
 DB 1159 CTTCAGAAAGCTGAGATATGAAATAATTAAGACATGATAACTGTCTCCACTATATGCT 1218
 OY 294 ValIleLeuHisGlnGluLysTyrIleValMetGlnHisValSerLysGluAsnLeuVal 313
 DB 1219 GTTGTCTTGAAACACCAATTTTACATTTGCTACAGTAATTTATGTCMAAGAGAACCTTATTA 1278
 OY 314 AspPheLeuArgThrArgGluArgAlaLeuValAsnThrAlaGluLeuGluIlePheSer 333
 DB 1279 GATTTCCTTAAGCAGAGAGATGGAATGATTTGCACTGCTCCACAGCTGTGATATGGCT 1338
 OY 334 LeuHisValAlaGluGluMetGluTyrLeuGluSerLysLysLeuValHisAspArgLeu 353
 DB 1339 GCTCAAGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1398
 OY 354 AlaAlaArgAsnIleLeuValSerGluAspLeuValAlaValSerAspPheGlyLeu 373
 DB 1399 CGGGCTGCTAAATATCTGTAGAGAAATCTGTGTGCAAAATACAGACCTTTGGTTTA 1458
 OY 374 AlaLysAla-----GluArgLysGlyLeuAspSerArgLeuProValLys 389
 DB 1459 GCAAGGTTAATTGACACAAATGATACACAGCAAGCAAGGTCGCAAAATTTCCAAATCAAA 1518
 OY 390 TrpThrAlaProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspValTrpSer 409
 DB 1519 TGGACAGCTCTGAACTGACAGCTGTATGCTGCTTACAAATTAAGTGTGATGCTGTGCTCA 1578
 OY 410 PheGlyValLeuLeuTrpGluValPheSerTyrGlyArgAlaProTyrProLysMetSer 429
 DB 1579 TTGGAAATCTGCAACACAGAACTAGTAACAAAGGGCGAGTGCATATCCAGATAGCGTG 1638
 OY 430 LeuLysGluValSerGluAlaValGluLysGlyTyrArgMetGluProProGluGlyCys 449
 DB 1639 AACCGTAGAGTACTAACAAGTGGAGCGAGATACAGATGCCCTGAGGCTGCT 1698
 OY 450 ProGluProValHisValLeuMetSerCysTrpGluAlaGluProAlaArgPro 469
 DB 1699 CCAAGAACTCCCTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1758
 OY 470 ProPheArgLysLeuAlaGluLysLeuAlaArgGluLeuArgSerAlaGluAlaProAla 489
 DB 1759 ACATTTGATATATTCAGCTCTTTTG----- 1785
 OY 490 SerValSerGlyLysAlaAspLysSerThrSerProArgSerGln 505
 DB 1786 -----GAAGACTACTTCACTGCTACAGAGCCACAGTACCAG 1821

RESULT 12
 US-10-175-523-50
 ? Sequence 50, Application US/10175523
 ? Publication No. US20030096264A1
 ? GENERAL INFORMATION:
 ? APPLICANT: Brockman, Jeffrey
 ? APPLICANT: Evans, David
 ? APPLICANT: Hook, Derek
 ? APPLICANT: Klimczak, Leszek
 ? APPLICANT: Laeng, Pascal
 ? APPLICANT: Palfreyman, Michael
 ? APPLICANT: Rajan, Prithi
 ? TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
 ? FILE REFERENCE: 3235/1J795-US3
 ? CURRENT APPLICATION NUMBER: US/10/175,523
 ? PRIOR FILING DATE: 2002-06-18
 ? PRIOR APPLICATION NUMBER: US 60/299,151
 ? PRIOR FILING DATE: 2001-06-18
 ? PRIOR APPLICATION NUMBER: US 60/317,828
 ? PRIOR FILING DATE: 2001-09-07
 ? PRIOR APPLICATION NUMBER: US 60/325,150
 ? PRIOR FILING DATE: 2001-09-25
 ? PRIOR APPLICATION NUMBER: US 60/333,047
 ? PRIOR FILING DATE: 2001-11-14
 ? PRIOR APPLICATION NUMBER: US 60/349,936
 ? PRIOR FILING DATE: 2002-01-18
 ? PRIOR APPLICATION NUMBER: US 60/361,834
 ? PRIOR FILING DATE: 2002-03-04
 ? NUMBER OF SEQ ID NOS: 197
 ? SOFTWARE: PatentIn version 3.1
 ? SEQ ID NO 50
 ? LENGTH: 2298
 ? TYPE: DNA
 ? ORGANISM: Homo sapiens
 ? US-10-175-523-50
 Alignment Scores:
 Pred. No.: 1,85e-63 Length: 2298
 Score: 709.50 Matches: 167
 Percent Similarity: 53.23% Conservative: 97
 Best Local Similarity: 33.67% Mismatches: 179
 Query Match: 26.56% Indels: 53
 DB: 14 Gaps: 14
 US-09-977-260-2 (1-507) x US-10-175-523-50 (1-2298)
 OY 35 ProProValSerAlaArgMetProThrArgArgTrp-----Ala 48
 DB 427 CGAGTCCAGAAATCTGACGCTTTTACCTGGACAGAGTTTCAAACTAAAGTCCAGAGAA 486
 OY 49 ProGlyThrGlnCysIleThrLysCysGluHisThrArgProLysProGlyGluLeuAla 68
 DB 487 CAAGGACATTTGTGACCTTTGTACCCCTATGATGATGATGATGATGATGATGATGATGAT 546
 OY 69 PheArgLysGluAspValValIleLeuGluAlaCysGluAsnLysSerTrpArg 88
 DB 547 TTCAGAAAGAGAGAGATGAAGATGAAAGTCTGGAG-----GACATGTGAGAAATGTGGAAA 600
 OY 89 ValLysHisThrSerGluGluGluGluGluLeuAlaGlyAlaLeuArgGluArg 108
 DB 601 GCAAACTCCCTTTTAAACAAAAAGAGGCTTCATCCCGAGAACTAGTGGCCAAATCTC 660
 OY 109 GluAlaLeuSerAlaAspProLysLeuSerLeuMetProTrpPheHisGlyLysIleSer 128
 DB 661 AACACCTTTGAAGAAAGAA-----GAGTGGTTTTTCAAGATATTAACC 702
 OY 129 GlyGlnGluAlaValGlnLeuGluIleProProGluAsp-----GlyLeuPheLeuVal 146
 DB 703 AGCAAGAGACGACAGAAAGCCTTTTGGACACAGCAAGAAATAGGCGTGGAGCTTTCTTAT 762
 OY 147 ArgGluSerAlaArgHisProGlyAspArgValIleCysValSer----- 161
 DB 763 AGAGAAAGTGAACATTTAAAGAGCTTCTCTGTCTGTGACAGACTTTGACCCCTGTG 822

Oy	162	Phehlgiaargvavaliilehstvyarxvaldeuhisargasp---glynhislenethrile	180
Db	823	CATGGTGTATGTTATTAAAGCACTACAAATTGAAAGTGGATTAATGGGGCTATTACATC	882
Oy	181	Aspeliualvalphepcysasnleuemetaspmetvalglnhistryserlysasplys	200
Db	883	TCTCCAGCAATCACTTTCCCTGTATGACGACGATGATTAAACATTACCAAAAGCAGGCA	942
Oy	201	GLyvallellecysthrlyleu-----valargprolys---arglyhisgely	215
Db	943	GATGGCTTGGTCCAGAAAGATTGGAGAGGCTTGATTGTATGTCCCAAGCCACAAAGCCATGG	1002
Oy	216	ThrylseseralagluglugluLeualaargalaclytrpleuleuAsnleuglnhisleu	235
Db	1003	GATTAAGATGCC-----TGGAGATGCCCGGGGAGTCCATC	1038
Oy	236	ThrlenglyalaglnleeglyglugluPheclayalavalLeuglnlelyglutryleu	255
Db	1039	AAGTTGGTGAAGAAAGCTGGCGCGGGAGTGGTGGGAAAGTGGATGGGTACTATPAC	1098
Oy	256	GlyGln---LysValAlaValalysasnleuys---cysasparvalthralaglnalphe	273
Db	1099	AACAGTACCAAGGTGGCTGTAAAAACCTGAAAGCCAGGAACTATGTCTGTCCAAAGCTTC	1158
Oy	274	LeuaspeluthrralavalamethcThrLysmetGlnhisglnasleuvalargleuengly	293
Db	1159	CTGGAAGAGCAACCTCATGAAACCCCTGCAGCATGACAGACCTGTCAGGCTCTACCT	1218
Oy	294	Vallelleu---HisglnlelyleutryllevalmetGlnhisvalserlysglyasleu	312
Db	1219	GTGGTCAACAGGAGGAGCCCATTTACTATCAACCGATCATGGCCAAAGGCAAGTTG	1278
Oy	313	ValasnPhelauArgThrArgGlyAargAlaLeuvalAsnThrAlaglnleuenglnphe	332
Db	1279	CTGATTTCCGTGAAGAGGATGAAGGTGGCAAAAGTGCCTTCCAAACCTGATTGACTTT	1338
Oy	333	SerleuhtisValAlaGluglymetGluTrrleugluseryLysLysleuValHisargasp	352
Db	1339	TCTCTCAGATTGGCAGAGGGAATGGCAATCACTCGAGCGGAAGAACTACATTACCCGGGAC	1398
Oy	353	LeuAlaAlaArgasnilleLeuValSerLysLysleuValAlaLysValSerAspPhegly	372
Db	1399	CTGGAGCAGCACTAATGTCTGTGCTCCAGTCACTAATGTGCCAAATTGCGAGATTTTGGC	1458
Oy	373	LeuAlaLysAla-----GluArglysglyLeuaspSerSerArgleuProval	388
Db	1459	CTTCTCAGTAAGTAATTGAAGATTAATGAGTACACAGCAAGGAAAGGCTAAACTTCCCTATT	1518
Oy	389	LysTrpThrAlaProGluAlaLeuLeuHisglnLysPheThrserLysSerAspValarp	408
Db	1519	AAGTGGAGGGCTCCGAGAACATCACTTTGGATGTTCCACTATTAACTCGATGTGTGG	1578
Oy	409	SerPheglyValleLeuTrrgluValPheSerTrrglyArgAlaProtyrProlysmet	428
Db	1579	TCTTTTGGAAATCTCTCTATACGAAATTGTCACTATGGGAAATAATCCCTACCAGGGAGA	1638
Oy	429	SerleuLysgluValSerLysAlaValaluglysglyTrrArgmetGluProProglucly	448
Db	1639	ACTATATGCCGAGCTGATACGCCCTGTGCCAGGGCTTCAGAGATCCCGGTGGAGAAC	1698
Oy	449	CysProGlyProValHisValleumetSerSerCysTrrgluAlagluProAlaArgarg	468
Db	1699	TGCCCAAGTGAAGCTCATGACATTATGAATAATGTCTGGAAAGAAAGCAGAAAGAGAGA	1758
Oy	469	ProProPheArgLysLeuAlaGluLysLeu-----	478
Db	1759	CCAAAGCTTGTACTATTACAGAGGCTCCTGGATGATTTCTACACAGCCAGGAAAGGCA	1818
Oy	479	---AlaArgGluLeuArgSerAlaglyLalProAlasertsergly	493
Db	1819	TACCAAGCAAGCACTTAAAGCAAGAGGAGACCCGTCCATTTTGGCAGGG	1866

```

RESULT 13
US-09-880-107-3710
; Sequence 3710, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3710
; LENGTH: 5527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 x16416
US-09-880-107-3710

Alignment Scores:
Pred. No.:      4,24e-62      Length:      5527
Score:          701.50       Matches:     178
Percent Similarity: 48.73%   Conservative: 90
Best Local Similarity: 32.36% Mismatches:    207
Query Match:     26.26%     Indels:       75
DB:              10         Gaps:        16

US-09-977-260-2 (1..507) x US-09-880-107-3710 (1..5527)

QY      2 AAGAGAATGAGTGGT-----LeuValSerTrpArg----- 11
      ||| ||||| ||||| ::::| |||||
DB      106 GGTGGGAAGAGGGGCTCCGGGCCGAGTCGTTCCGTGGCGCAAAATGTTGGAGATCGCCTG 165
QY      12 AAlarPniisgLyCyuaSPserdlaagLugluLeuProalrvalSerProalgrPheLuarg 31
      ||||| ||| |::: ||| ||| |||
DB      166 AAGCTGTGGTGCGCTGCACAAATCCAAAGAGGGGCTGCTCCGTCSCAGCTGTATTCTGGAA 225
QY      32 AAtGrPiHSrProrProToValSerdlAaayMeProrThlArg----- 45
      ||||| ::::| :|:::|
DB      226 GAAGCCCTTCACGGCGCATGACATCTGACTTGAAGCCCTCAGGGCTCGAGTAAGCCCGCT 285
QY      46 ATGTGTAlArProglYThrgLnCyslIethrLYscYglnInshIthrArProLYS----- 63
      ||||| ::: |||
DB      286 CGTTGGAACTCCAGAGAAACCTTCTCGGTGGACCCGAAAGAAATGACCACACCTTTTC 345
QY      64 -----ProglYclLleuAlArPheargLYsgLYaSP 73
      ||:::| ||||| ::::|
DB      346 GTTGCACSTGTATGATTTTGTGGCAGCATGGAGATMACACTCTGAACTAACATAAGGTGAA 405
QY      74 ValValThrlleLengluLaCysguLaaNlyssertTrprYalrYalulYsnInshIstr 93
      ::::| ||| ||| ||| ||| |||
DB      406 AAGCTCCGGGTGTGGTAGTATATATCAACAATGGGAATGGnTGAAGCCCCA-----ACC 459
QY      94 SerGlYglnGLylLeuLeuAlaAlaGLyAlaLeuAlrGLyuArGLyAlaLeuSerAla 113
      ::::| ::: ::: |||
DB      460 AAAAATGGCAAGCGCTGGCTCCCAACCACTACAGCACGCAAGCAAGTGTGGAGAA 519
QY      114 AsPrroLYslLeuSerLeuMetProrThrPheInshgLYslIesetGLyngluAlaVal 133
      ||::::| ||||| ::::| |||
DB      520 CAC-----TCCtGGTGCATATGGGCGCTGTGTCCCGCAATGGCGGTGAG 561
QY      134 gInglInLeuGlProrToGlnsPrGLyLeuPrleuValArgGuLYserdlAaayInshIstr 153
      ||| ::::| ||||| ||||| |||
DB      562 TATCTGCTGAGCAGCGGAGTCAATGGCACTTCTTGCTGAGTGAAGTGAAGACAGTCTCT 621

```

QY 154 GLYASPTyValLeucysValSerPheGlyArgAspValIleHisTyrArgVal---Leu 172
||| : : : : : ||| ||||| : : : : :
Db 622 GGCCAGAGGTCCATCTCGGTGAGATGAGAGGAGGAGGTGTATACAGACATCAACACT 681
QY 173 HisArgAspGlyHisLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspMet 192
||| ||| : : : : : ||| ||| : : : : :
Db 682 GCCTTCGATGCGCAAGCTCTACGCTCTCTCCGAGACGCCCTTCAACACCCCTGGCCGAGTTG 741
QY 193 ValGlnHisTyrSerLysAspLysGlyAlaIleCysThrLysLeu-----ValArgPro 210
||| : : : : : : : : : : : : : : : |||
Db 742 GTTCATCATCATTCATCAACGCGTGGCCGAGGCTCATCACCCGCTCATTTATTCAGGCCCA 801
QY 211 LysArgLys-----HisGlyThrLysSerAlaGluIleGluAlaArgAla 226
||| : : : : : : : : : : : : : : :
Db 802 AACGGCAACAGCCCACTGTATAGGTGTGTCCCACTACACACAG----- 849
QY 227 GlyTyrLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnIleGlyGluIlePhe 246
||| : : : : : : : : : : : : : : :
Db 850 ---TGGGAGATGAGACCGCACGACATCACCATGACACACAGCTGGCGGGGCGGCGCTAC 906
QY 247 GlyAlaValLeuGlnGlyGlyTyr-----LeuGlyGlnLysValAlaValLysAsnIle 264
||| : : : : : : : : : : : : : : :
Db 907 GGGGAGGTGTACAGCGGCGCTGTGAGAAATACAGCCTGACGCTGGCGCGCTGAGAGCCTTG 966
QY 265 LysCysAsp---ValThrAlaGlnAlaPheLeuAspGluThrAlaValMetThrLysMet 283
||| : : : : : : : : : : : : : : :
Db 967 AAGGAGACACCCATGGAGGTGAGTGTGAAAGCTGTGAAAGCTGACCTGATGAAAGAGATC 1026
QY 284 GlnHisGluAsnLeuValArgLeuGlyValIleLeuHisGln---GlyLeuTyrIle 302
||| : : : : : : : : : : : : : : :
Db 1027 AAGACCCCTAACCTGGTGCAGCTCTTGGGCTGTGACCCGGAGGCCCGCTTCTATATC 1086
QY 303 ValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgAla 322
||| : : : : : : : : : : : : : : :
Db 1087 ATCAGTAGTGTCTGACCTACAGGGAACCTCTGACTGAGGAGAGTGCACACCGGAC 1146
QY 323 LeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGluIleMetGlyTyr 342
||| : : : : : : : : : : : : : : :
Db 1147 GAGGTAGACGCCGTGTCTGTCTGATGAGCCACATCTCTGACAGCCATGAGAGTAC 1206
QY 343 LeuGlnSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuValSerGlu 362
||| ||| : : : : : : : : : : : : : : :
Db 1207 CTGGAGAGAGAAACTTCATCCACAGATCTTGTGCTCCGAAAGTCCCTGTGATGGGAG 1266
QY 363 AspLeuValAlaLysValSerAspPheGlyLeuAlaLys----- 375
||| : : : : : : : : : : : : : : :
Db 1267 AACCACTGTGTAAGGTAGCTGATTTGGCTGACAGAGTGTGACAGGGACACTTAC 1326
QY 376 AlaGluArgLysLysAspSerArgLeuProValLysThrPheAlaProGluAla 395
||| : : : : : : : : : : : : : : :
Db 1327 AACGCCCATGCTGA-----GCCAAGTCCCATCAATGAGTGCACCCGAGAGC 1377
QY 396 LeuLysHisGlyLysPheThrSerLysSerAspValTyrSerPheIleValLeuLysTyr 415
||| : : : : : : : : : : : : : : :
Db 1378 CTGGGCTTACACAGATCTCTCCATCAAGTCCGAGCTGTGGCATTTGGAGATTTGCTTGG 1437
QY 416 GluValPheSerTyrGlyArgAlaProTyrProLysMetSerLeuLysGluValSerGlu 435
||| : : : : : : : : : : : : : : :
Db 1438 GAATTTGCTACCTATGCGATGTGCCCTTACCCGGAATGATGACTGTCCAGAGTGTATGAG 1497
QY 436 AlaValGluLysGlyTyrArgMetGluProProGluIleCysProGlyProValHisVal 455
||| : : : : : : : : : : : : : : :
Db 1498 CTGCTAGAGAGAGACTATCCGATGAGGCCCAAGAGGTGCCCCAAGAGAGCTATGAA 1557
QY 456 LeuMetSerSerCysTyrPheLysArgLysProAlaArgArgProPheArgLys----- 473
||| : : : : : : : : : : : : : : :
Db 1558 CTTCATCGAGCATGTTGGCAGTGGAAATCCCTCTGACCGGCGCTCTTGTCTGAATCCAC 1617
QY 474 -----LeuAlaGluLysLeuAlaArgGluLeu 482
||| : : : : : : : : : : : : : : :
Db 1618 CAAGCTTTGAAACAATGTTCCAGAGATCCAGATATCTACAGAGAACTGGAAAGAGAGCTG 1677
QY 483 ArgSerAlaGlyAlaProAlaSerValSer 492

Db 1678 GGGAACAAAGCGCTCCGTGGGCGCTGTGAGT 1707
RESULT 14
US-09-967-768A-300
; Sequence 300, Application US/09967768A
; Patent No. US2002015087A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 300
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-300
Alignment Scores:
Pred. No.: 1 52e-60 Length: 2354
Score: 681.50 Matches: 176
Percent Similarity: 49.03% Conservative: 78
Best Local Similarity: 33.98% Mismatches: 198
Query Match: 25.51% Indels: 67
DB: Gaps: 13
US-09-977-260-2 (1-507) x US-09-967-768A-300 (1-2354)
QY 23 ProArgValSerProArgPheLeuArgAlaTyr----- 33
||| : : : : : ||| : : : : :
Db 185 CCACGG-----CCAAAGAGAGATGTGGCTGGAAGGGAGCTTCAGAGCTACGGGGAC 238
QY 34 -----HisPro-----Pro 36
||| : : : : : ||| : : : : :
Db 239 CAGACCAATAGGGCTGACCCCAATAGGCCCGGCTGATCTATTTGCCACATCC 298
||| : : : : : ||| : : : : :
QY 37 ProValSerAlaArgMetProThrArgArgTyrAlaProGly----- 50
||| : : : : : ||| : : : : :
Db 299 CCAACTACAGCACTTCTCTCTCAGGCCATC-AACCTGTGCTTCTTGTAGTGGCACC 357
||| : : : : : ||| : : : : :
QY 51 -----ThrGlnCysIleThrLysCysGlnHisThrArg 61
||| : : : : : ||| : : : : :
Db 358 ATCAGGGGTGTCTCAGGAGATTTGGGTGACCCCTGTCTCATCTGATGACTATGAGGT 417
||| : : : : : ||| : : : : :
QY 62 ProLysProGlyLysLeuAlaPheArgLysGlyAspValValThrIleLeuGluAlaCys 81
||| : : : : : ||| : : : : :
Db 418 CGAACTGAGAGATCTCACTCACTTCCACCAAGGGAGAGTTCACCACTCTGAAACAATACT 477
||| : : : : : ||| : : : : :
QY 82 GluAsnLysSerTyrTyrArgValLysHisIleThrSerGlyGlnGlyLeuAla 101
||| : : : : : ||| : : : : :
Db 478 GAAAGT---GACTGGTGGAGAGCTCGCTCTCAGCTCCGGAAAGATGGTGGCATTTCC 534
||| : : : : : ||| : : : : :
QY 102 AlaGlyAlaLeuArgGluArgLysAlaLeuSerAlaAspProLysLeuSerLeuMetPro 121
||| : : : : : ||| : : : : :
Db 535 AGCAACTAGCTGGCCCTGTGTACCTGATCAATCAAGCTGAA-----GAG 576
||| : : : : : ||| : : : : :
QY 122 TyrPheHisGlyLysIleSerGlyGlnAlaValGlnGlnLeuGlnProPro----- 139
||| : : : : : ||| : : : : :
Db 577 TGTACTTTTGGAAGATTTGGAGAAAGATGACAGAGAGCAGCTGCTTTCACACAGGCAAC 636
||| : : : : : ||| : : : : :
QY 140 GluAspGlyLeuPheLeuValArgLysSerAlaArgHisProGlyLysTyrValLeuCys 159
||| : : : : : ||| : : : : :
Db 637 CCCAGGGGCGCTTCTCATTTGCGGAAAGCGAGACCCACCAAGGTGCTACTTCCTGTCC 696

```

QY 160 Val-----SerPheGlyArgAspValIleHisTyrArgValLeuHisArg 174
Db 697 ATCCGGAGCTGGATCGACGACGAGGCGATCATGTGAAGCATTTCAAGATCCGAAATCG 756
QY 175 Asp---GlyHisLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspMetVal 193
Db 757 GACATGGCGCGCTACTACATCACCACACAGGGGTTTCAGTTCAACATGGTGACAGGAGCTGTG 816
QY 194 GluHisTyrSerLysAspLysGlyAlaIleCysThrLysLeuValAArgProLysArgLys 213
Db 817 CAGACACTACATGAGGTGAATGACGGCTGTGCAACCTGTCTCAATCGCGCTGCACCATC 876
QY 214 HisGlyThrLysSerAlaGluGluGluAlaValArgAlaGluThrLeuAsnLeuGln 233
Db 877 -----ATGAAGCCGCGAGAGCTGGCTGGCTGCCAAGACGCTGGAGATCAGCCGAC 930
QY 234 HisLeuThrLeuGlyAlaGlnIleGlyGluGluGluPheGlyAlaValLeuGlnGlyGlu 253
Db 931 TCCATCAGCTGGAGCGCGCTGGCGACCGGCTGTCTGGGGATGTGTGGCTGGGACG 990
QY 254 TyrLeuGlyGln---LysValAlaValLysAsnLys---CysAspValThrAlaGln 271
Db 991 TGGAAACGGCAGCACTAAGGTGGCGGTGAAGCGCTGAAGCGCGGACCATGTCCCGCAG 1050
QY 272 AlaPheLeuAspGluThrAlaValMetThrLysMetGlnHisGluAsnLeuValArgLeu 291
Db 1051 GCCTCTCTGGAGAGGCGCGAGGTATGACGTGCTGGCGACGCAAGCATGTGGTCACTG 1110
QY 292 LeuGlyValIleLeuHisGlnGlyLeuTyrIleValMetGlnHisValSerLysGlyAsn 311
Db 1111 TACCGCTGGTGTGCGAGAGGCCATCTACATCGTGAACGAGTGTGTGACGAGC 1170
QY 312 LeuValAsnPheLeuAlaGlyThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuGln 331
Db 1171 TTCTGGATTTTCTCAAGAACCCAGAGCGCGAGATTGTAGCGCTCCCAATTGGTGAC 1230
QY 332 PheSerLeuHisValAlaGluGlyMetGluTyrLeuGluSerLysLysLeuValHisArg 351
Db 1231 ATGGCAGCCAGCAAGTGAAGGCGATGCGCTACATCGAAGCGATACATTCACCGC 1290
QY 352 AspLeuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaLysValSerAspPhe 371
Db 1291 GACCTGAGGCGCAACATCTGTGGGAGCGCTGTGCGTGAAGATCGCAGACTTT 1350
QY 372 GlyLeuAlaLysAla-----GluArgLysGlyLeuAspSerSerArgLeuPro 387
Db 1351 GCGTTGGCGCTCATCAAGAGAGATGATGACAAACCTGCCAAGGTTCCAACTTCCC 1410
QY 388 ValIleTyrThrAlaProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspVal 407
Db 1411 ATCAAGTGAACAGCCCAAGAGCTGCTTTGGCAGATTCCACCATCAAGTCAAGAGCTG 1470
QY 408 TrpSerPheGlyValLeuLeuTyrPgluValPheSerTyrGlyArgAlaProTyrProLys 427
Db 1471 TGGCTCTTGGAGCTGCTGCTACAGCTCATCAAGAGCGCAAGGCTCCCTATCCAGGC 1530
QY 428 MetSerLeuLysGlyValSerGluAlaValGluLysGlyTyrArgMetGluProGlu 447
Db 1531 ATGATATAAGGGAAGTGTGGAACAGTGAAGAGGAGTACACATGCGCTGCCCTCA 1590
QY 448 GlyCysProGlyProValHisValLeuMetSerSerSerSerSerSerSerSerSerSer 467
Db 1591 GGGCGCCAGCATCCCTGTACAGAGCATGAACAGACAGCTGCGCTGGACCGGAGAG 1650
QY 468 ArgProPheArgLysLeuAlaGluLysLeuAlaArgLysLeuArgSerAla 485
Db 1651 AGGCTACTCTGAGTACTGACTGCTCTCTGAGAGACTACTTCACTCCGCT 1704

```

RESULT 15
 US-09-977-269-5
 Sequence 5, Application US/09977269
 Patent No. US20020082037A1

```

: GENERAL INFORMATION:
: APPLICANT: ULIRICH, AXEL
: APPLICANT: GISHIZKY, MIKHAIL
: APPLICANT: SURES, IRMINARD
: TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
: FILE REFERENCE: 038602/1260
: CURRENT APPLICATION NUMBER: US/09/977, 269
: PRIOR FILING DATE: 2001-10-16
: PRIOR APPLICATION NUMBER: 08/232, 545
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 5
: LENGTH: 2770
: TYPE: DNA
: ORGANISM: Unknown Organism
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (366)..(1880)
: FEATURE:
: OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
: OTHER INFORMATION: Kinase 3
: US-09-977-269-5

Alignment Scores:
Score: 1,2e-57 Length: 2770
Percent Similarity: 654.50 Matches: 153
Best Local Similarity: 53.068 Conservative: 81
Query Match: 34.698 Mismatches: 166
DB: 24.508 Indels: 41
Caps: 12

US-09-977-260-2 (1-507) x US-09-977-269-5 (1-2770)
QY 66 GluLeuAlaPheArgLysGlyAspValValThrIleLeuGluAlaCysGluAsnLysSer 85
Db 543 GACTTGCAGCTTCCAGCAGCGTGAACAACTTCAAGTCTGGAACCTTGG---CATGAGCGC 599
QY 86 TrpIleArgValLysHis-----HisThrSerGlyGlnGluGly 98
Db 600 TGGTGGTTTGGCAGACACTTGGAGAAAGACAGATGGCTCCAGTCAAGCAATCAAGC 659
QY 99 LeuLeuAlaAlaGluAlaLeuArgLysGluAlaLeuSerAlaAspProLysLeuSer 118
Db 660 TATATTCCTTCTAATCACTGAGCGGTGAGAGACAGAAAGCTTACAGCGAGAG----- 707
QY 119 LeuMetProTyrPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnPro 138
Db 708 -----CCGTGGTCTTGTGAGCAATCGAAGATCAATGATCGAGAAACAACTATTAT 761
QY 139 ProGluAsp-----GlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyr 156
Db 762 TCAGAAAACAGACCGGCTTCTCTTCTAATCAAGAAAGTGAAGAAAGCAAGAAAGGAAATTC 821
QY 157 ValLeuCysValSerPheGlyArgAspValIleHisTyrArgValLeuHisArgAsp--- 175
Db 822 TCTCTTCAAGTTTAAGTAGGAGCAGCTGTAAACAGTAAAGAAATTAAGAGCTGATGA 881
QY 176 GlyHisLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspMetValGluHis 195
Db 882 GGGGGATTTTTCACAGCGAGAAAGATCTTTCAACACTGAAAGCAAGAAATTTGTGAGCCAC 941
QY 196 TyrSerLysAspLysGlyAlaIleCysThrLysLeuValAArgProLysArgLys----- 213
Db 942 TACACCAAGACAACTGACGCGCTGTGTGTCAAGCTGGGAAACCATGCTTAAAGATCCAG 1001
QY 214 -----HisGlyThrLysSerAlaGluGluGluAlaValArgAlaGly 227
Db 1002 GTCCAGCTGCATTTGATTTGTGCTATAAAGCGTGACCA----- 1043
QY 228 TrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnIleGlyGluGluPheGly 247
Db 1044 TGGGAGATAGACCGCAACTCATCAAGCTTCTGAAAGCGAATGGAGTGTGGTGTGGC 1103

```

```
QY 248 AlaValLeuGlnGlyTyrLeuGly---GlnLysValAlaValLysAsnIleLys--- 265
Db 1104 GAAGTATGGGAGAGTGTGTGGAAACAATACCACATCCAGTACAGTGAACATTTAAACCA 1163
QY 266 CysAspValThrAlaIleAspLeuAspLeuThrAlaValMetThrLysMetGlnHis 285
Db 1164 GGTTCATGAGTCCAAATGACTTCTGAGGAGGACACAGATTAATGAAGAACTTAAGCAT 1223
QY 286 GluAsnLeuValArgLeuGlyVal---IleLeuHisGlnGlyLeuTyrIleValMet 304
Db 1224 CCAAGCTTATCCAGCTTTTGTGCTTTGACCTTTGAGACATCCAAATTTATATTATTA 1283
QY 305 GlnHisValSerLysGlnLeuValAsnPheLeuArgThrArgLysArgAlaLeuVal 324
Db 1284 GAGTTGATGACATGGAAGTCTGCAAGAATATCTCCAAATATGACACATGATCAAAAATC 1343
QY 325 AsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGlnGlyMetGlyTyrLeuGlu 344
Db 1344 CATCTAGCTCAACAGGTAGACATGGCGGCACAGGTTGCCCTGGAATGGCCTATCTGGAG 1403
QY 345 SerLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuValSerGluAspLeu 364
Db 1404 TCTCGGAACTACATTCACAGAGATCTGGCTGCCAGAAATGCTCTGTTGGTGAACATAT 1463
QY 365 ValAlaLysValSerAspPheGlyLeuAla-----LysAlaGlnArgLysGlyLeu 381
Db 1464 ATCTACAAAGTAGACATTTTGGACTTGGCCAGAGTTTAAAGTAGATATATGAAGACATC 1523
QY 382 AspSerSerArg-----LeuProValLysTrpThrAlaProGlnAlaLeuLys 397
Db 1524 TATGAATCTAGACACGAAATAAAGCTGCCGGTAGAGTGGACTGGCCGGAAGCCATTCGT 1583
QY 398 HisGlyLysPheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluVal 417
Db 1584 AGTAATAAATTCAGCATTAAGTCCGATGTATGCTCATTTGGAAATCCTTCTTTATGAATC 1643
QY 418 PheSerTyrGlyArgAlaProLysMetSerLeuLysGlnValSerGluAlaVal 437
Db 1644 ATTACTTATGGCAAAATGCTTACAGTGTATGACAGGTGCCAGGTAATCCAGATGTTG 1703
QY 438 GlnLysGlyTyrArgMetGluProProGlnGlyCysProGlyProValHisValLeuMet 457
Db 1704 GCTCAAACTATAGACTTCGGCAACCATCCAACTGTCACAGCAATTTTACACATCATG 1763
QY 458 SerSerCysTrpGlnAlaGluProAlaArgArgProPheArgLysLeuAlaGluLys 477
Db 1764 TTGGAGTGGTGAATGACAGAGCCTTAAGGAACGACTACATTTGAGACACTGCTGGAAA 1823
QY 478 Leu 478
Db 1824 CTT 1826
```

Search completed: August 2, 2003, 05:12:26
Job time : 1007 secs